SIMAI e-Lecture Notes
ISSN 1970-4429

Mathematical Theory of Electromagnetism

Piero Bassanini

University of Rome, "La Sapienza" Rome, Italy
Piero.Bassanini@uniroma1.it

Alan Elcrat

Wichita State University, Wichita, Kansas
elcrat@math.wichita.edu

DOI: 10.1685/SELN09001 - Vol. 2 - 2009

Contents

List of Figures

Preface

These lecture notes treat the mathematical theory of electromagnetism. They are written at a level appropriate for graduate students in mathematics. Very little is assumed about prior exposure to the physical theory, but some mathematical sophistication is assumed at various points in the exposition. The subject is treated as a continuum theory with only brief mention of underlying molecular origins of phenomena. Also, thermodynamical considerations are not emphasized.

The basic organization of these lectures was made by the first author and used in lectures given at Universita di Roma, La Sapienza. The purpose of them is to give a self contained treatment of electromagnetism for students of mathematics in order to lay the foundation for the many applications of these ideas in applied mathematics. In earlier times it might have been assumed that students of mathematics that had use for this material would take courses in physics or electrical engineering departments, but the crowded curriculum and prerequisite structure make this difficult today. In addition an exposition written from the point of view of applied mathematics is natural for students of mathematics. This applies in particular to the treatment of steady currents in chapter 3 and of nonlinear optics in chapter 6. For these reasons lecture notes of the kind presented here can play an important role for students of mathematics.

Chapter 1

Maxwell Equations For Bodies At Rest

An exposition of the Maxwell theory of electromagnetism must perforce begin by stating the system of physical units chosen, as the Maxwell equations change (albeit slightly) according to this choice. The basic well-known experimental fact underlying the theory is the existence of the electric charge, which is convenient to view as an autonomous physical quantity endowed with an independent physical dimension. As unit charge we might think in principle of taking the charge of the electron, an elementary particle carrying a negative electric charge $-e$. This electron charge, though, is so small that it turns out to be impractical in view of describing electromagnetic phenomena at a macroscopic level, which is the level of description targeted in Maxwell's model. A convenient practical charge unit is the coulomb , chosen so that the charge of the electron is given in absolute value by

$$e = 1.6\,10^{-19}\,coulomb$$

The related electric current unit is the ampere , a practical unit defined by the ratio coulomb/second. We are thus led to adopt the MKSA system of units, based on the choice of four fundamental quantities and four corresponding independent units (see the Appendix at the end of the chapter).

1.1 Preliminaries and basic experimental facts

1.1.1 Mathematical Preliminaries

We will call domain any open connected set in $\mathbb{R}^N$. If Ω is a domain, we will denote by $\partial\Omega$ its boundary, by $\overline{\Omega} = \Omega \cup \partial\Omega$ its closure, and by $\boldsymbol{n}$ the unit normal to $\partial\Omega$, oriented as a rule towards the exterior of Ω. We will use boldface letters for vectors, a dot for the scalar product in $\mathbb{R}^N$. The cross product of two vectors in $\mathbb{R}^3$ will be denoted by the symbol $\wedge$ used for exterior products in differential geometry. If $\boldsymbol{x}$,$\boldsymbol{y}$ are any two vectors in $\mathbb{R}^3$ then

$$\boldsymbol{x} = (x_1, x_2, x_3) = \sum_{k=1}^{3} x_k \mathbf{c}_k \qquad \boldsymbol{y} = (y_1, y_2, y_3) = \sum_{k=1}^{3} y_k \mathbf{c}_k$$

where $\mathbf{c}_k$ are the unit vectors of a cartesian reference frame, $\boldsymbol{x}\cdot\boldsymbol{y} = x_1y_1 + x_2y_2 + x_3y_3$, and $|\boldsymbol{x}|=\sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ is the modulus of $\boldsymbol{x}$. Domains in $\mathbb{R}^3$ will also sometimes be called "volumes". We will denote by $C^k(\Omega)$, $C^k(\overline{\Omega})$ the class of scalar or vector functions continuous (for $k = 0$) together with all partial derivatives up to the order k (for $k = 1, 2, ...$) in Ω or $\overline{\Omega}$, respectively.

A function will be called biregular in a domain Ω if it is of class $C^2(\Omega) \cap C^1(\overline{\Omega})$.

If Ω is a bounded domain in $\mathbb{R}^3$ with a smooth boundary $\partial\Omega$, then the Gauss Lemma

$$\text{(LG)} \qquad \int_\Omega u_{x_k} \mathbf{dx} = \int_{\partial\Omega} u n_k \, dS \qquad \left(u_{x_k} := \frac{\partial u}{\partial x_k},\ \ k = 1, 2, 3\right)$$

holds for all (one-valued) functions $u \in C^1(\overline{\Omega})$, with $\boldsymbol{n}=\boldsymbol{n}(\boldsymbol{x}) = (n_1, n_2, n_3)$ the unit exterior normal to $\partial\Omega$ at the point $\boldsymbol{x}= (x_1, x_2, x_3)$, $\boldsymbol{dx}= dV = dx_1dx_2dx_3$ the volume element in $\mathbb{R}^3$ and $dS = dS_x$ the surface element on $\partial\Omega$. In two dimensions, (LG) holds for Ω a bounded domain in $\mathbb{R}^2$, $\boldsymbol{dx}= dx_1dx_2$ the element of area and $dS = ds$ the arc length element along the curve $\Gamma = \partial\Omega$ (see Exercise 1).

The Gauss Lemma implies the divergence theorem

$$\text{(DT)} \qquad \int_\Omega \operatorname{div} \boldsymbol{v} dV = \int_{\partial\Omega} \boldsymbol{v} \cdot \boldsymbol{n} \, dS$$

which holds for any (one-valued) vector field $\boldsymbol{v} \in C^1(\overline{\Omega})$. (These regularity assumptions can be somewhat relaxed.) Another consequence of the Gauss Lemma is the identity

$$\int_{\partial\Omega} \boldsymbol{n} \wedge \boldsymbol{v}\, dS = \int_{\Omega} curl \boldsymbol{v}\, dx$$

which implies that the integral of $\boldsymbol{v} \wedge \boldsymbol{n}$ on $\partial\Omega$ (called vector circulation in aerodynamics) is zero if $\boldsymbol{v}$ is irrotational in the normal domain Ω. This fact also follows from the identity (Exercise 2)

$$\text{(VT)} \qquad \int_{\partial\Omega} \boldsymbol{n} \wedge \boldsymbol{v}\, dS = \int_{\partial\Omega} \boldsymbol{x}\, \boldsymbol{n} \cdot curl\, \boldsymbol{v}\, dS$$

which shows that the vector circulation depends in reality only on the normal trace of $curl\, \boldsymbol{v}$.

A "manifold" (curve or surface) will be called closed if it has no boundary points, otherwise it will be called open. This terminology is topologically consistent if we think of manifolds consisting only of interior points. For example, the boundary $\partial\Omega$ of a domain Ω in $\mathbb{R}^3$ is a closed surface (an open/closed set without boundary) and the boundary ∂S of an open surface S in $\mathbb{R}^2$ or $\mathbb{R}^3$ is a closed curve. We will sometimes write

$$\oint_S$$

to denote surface or contour integration over a closed manifold S (curve or surface).

A surface is called orientable if the normal $\boldsymbol{n}$ exists (almost) everywhere [19]. All surfaces considered in these notes are assumed once and for all to be orientable, so we will omit this specification further on.

If f, $\boldsymbol{v}$ are (one-valued) C^1 scalar and vector functions and S is a smooth open surface with unit normal $\boldsymbol{n}$, the two variants of the Stokes theorem hold:

$$\text{(ST1)} \qquad \int_S \boldsymbol{n} \cdot curl\, \boldsymbol{v}\; dS = \int_{\partial S} \boldsymbol{t} \cdot \boldsymbol{v}\, dS$$

$$\text{(ST2)} \qquad \int_S \boldsymbol{n} \wedge grad\, f dS = \int_{\partial S} f\, \boldsymbol{t}\; dS$$

where $\boldsymbol{t}$ is the unit tangent to the curve ∂S, oriented according to the right-handed screw rule with respect to $\boldsymbol{n}$. If S is closed ($\partial S = \emptyset$) the second members of (ST1) and (ST2) vanish, and for $f, \boldsymbol{v}$ of class C^2 the Stokes theorem follows from the Gauss Lemma and the identities $curl grad \equiv div\, curl \equiv 0$. Note that $\boldsymbol{n}\cdot curl$ and $\boldsymbol{n}\wedge grad$ are interior (tangential) differential operators on the "tangent bundle" of S, and $\boldsymbol{n}\cdot curl(\boldsymbol{n}\wedge grad)$ is the Laplace-Beltrami operator over S [19], which by force of (ST1) and (ST2) satisfies the relation

$$\int_S \boldsymbol{n} \cdot curl\,(\boldsymbol{n} \wedge gradf)\; dS \;=\; \int_{\partial S} \boldsymbol{t} \wedge \boldsymbol{n} \cdot gradf\, dS$$

A domain will be called normal if it is bounded and is such that the Gauss Lemma holds. For this it is sufficient that $\partial\Omega$ be a piecewise C^1 manifold [19].

In three dimensions the concept of a simply connected domain, well-known from Calculus for plane sets, can be defined in two different ways.

(i) A domain Ω in $\mathbb{R}^3$ is said to be contourwise simply connected if (like in 2D) every closed regular curve Γ contained in Ω is boundary of an open surface S entirely contained in Ω; in formulas, if

$$\Gamma \subset \Omega, \quad \partial\Gamma = \emptyset \quad \Rightarrow \quad \Gamma = \partial S\ , \quad S \subset \Omega$$

In a c.s.c. set every irrotational C^1 vector field can be written as the gradient of a global one-valued C^2 scalar potential:

$$curl\boldsymbol{v} = \boldsymbol{0} \ \text{ in } \Omega \ \text{ c.s.c.} \ \Rightarrow \ \boldsymbol{v} = -grad\, u\ , \ u \text{ one-valued in } \Omega$$

If this is not true Ω is said to be contourwise multiply connected.If the domain is c.m.c., the scalar potential u is in general many-valued in Ω. In this case, the Gauss Lemma is in general false unless a suitable "branch cut" or "branch surface" is introduced to make the domain simply connected (Exercise 3).

(ii) A domain Ω in $\mathbb{R}^3$ is said to be surfacewise simply connected if every closed regular surface S contained in Ω is boundary of a domain D entirely contained in Ω; in formulas, if

$$S \subset \Omega\ , \quad \partial S = \emptyset \quad \Rightarrow \quad S = \partial D\ , \quad D \subset \Omega$$

In a s.s.c. set a C^1 solenoidal (divergence-free) vector field can be written as the curl of a C^2 global one-valued vector potential:

$$div\boldsymbol{v} = \boldsymbol{0} \ \text{ in } \Omega \ \text{ s.s.c.} \ \Rightarrow \ \boldsymbol{v} = curl\,\boldsymbol{V}\ , \ \boldsymbol{V} \text{ one-valued in } \Omega$$

If this is not true Ω is said to be surfacewise multiply connected. If the domain Ω is s.m.c., $\partial\Omega$ is not connected and a (one-valued) vector potential $\boldsymbol{V}$ can be defined only locally in Ω [1]

A domain which is both c.s.c. and s.s.c. will be called simply connected. Its boundary is necessarily connected. As an example, a sphere is simply connected; a torus is simply connected in the sense (ii) but not in the sense (i); in contrast, the domain between two concentric spheres is simply connected in the sense (i) but not in the sense (ii).

We recall the meaning of the symbols

$$\begin{aligned} f = O(g) \quad &\Rightarrow \quad |f/g| \ \text{ remains bounded} \\ f = o(g) \quad &\Rightarrow \quad f/g \ \text{ tends to zero} \\ f \sim g \quad &\Rightarrow \quad f/g \ \text{ tends to } 1 \end{aligned}$$

in some limit operation (explicit or implicit).

If $\boldsymbol{v}(\boldsymbol{x})$ is a vector field, the lines of flow (or streamlines, or lines of force) of $\boldsymbol{v}$ are the integral curves of $\boldsymbol{v}$, i.e. the curves given in parametric form by $\boldsymbol{x}=\boldsymbol{x}(\tau)$ with

$$\frac{d\boldsymbol{x}}{d\tau} = \boldsymbol{v}(\boldsymbol{x}(\tau))$$

If v_1, v_2, v_3 do not vanish at the point $\boldsymbol{x}(\tau)$ these equations can be written as

$$\frac{dx_1}{v_1} = \frac{dx_2}{v_2} = \frac{dx_3}{v_3}$$

and $v_j = 0$ implies $dx_j = 0$, $j = 1, 2, 3$. The points where $\boldsymbol{v} = \boldsymbol{0}$ are singular points of the lines of flow, i.e. points where the tangent vector does not exist. Some results on the geometry of lines of flow are derived in Exercise 4.

Finally, we recall that the support of a function is the closure of the set where the function does not vanish.

To avoid burdening the exposition, mathematical details of a purely formal nature (like regularity assumptions, when they are obvious) will often be omitted and their completion will be left to the reader.

---

[1] in other words, a solenoidal vector field in a generic subdomain Ω of $\mathbb{R}^3$ cannot in general be written as a *curl* in all of Ω.

1.1.2 Electric charge and electric field. Displacement vector.

An electric charge Q can be concentrated at a point (point source) or distributed over surfaces or "volumes". In the presence of an electric field a charge undergoes a force, related to the electric intensity vector $\boldsymbol{E}$, defined as the mechanical force $\boldsymbol{F}$ acting over a unit pointlike test charge (a more precise definition is given in Exercise 5). The force acting over a charge q concentrated at a point $\boldsymbol{x}$ of any material medium at time t is then

$$\boldsymbol{F} = q\boldsymbol{E} \equiv q\boldsymbol{E}(\boldsymbol{x}, t) \tag{1.1}$$

Consider two points $\boldsymbol{x}_1$,$\boldsymbol{x}_2$ and a smooth path $\Gamma = \Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2)$ connecting them. Let s denote the curvilinear coordinate and $\boldsymbol{t}= d\boldsymbol{x}/ds$ the unit tangent along Γ. The mechanical work done by the electric field $\boldsymbol{E}$ along Γ, defined by the circulation of $\boldsymbol{E}$

$$V = \int_{\Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2)} \boldsymbol{E} \cdot \mathbf{t}\, ds \equiv \int_{\Gamma(\mathbf{x}_1,\boldsymbol{x}_2)} \sum_{k=1}^{3} E_k\, dx_k \tag{1.2}$$

is called electromotive force (e.m.f.). If the field $\boldsymbol{E}$ is irrotational in some domain Ω , that is if

$$\text{curl}\boldsymbol{E} = \mathbf{0}$$

in Ω , then there exists an electric potential u, such that

$$\boldsymbol{E} = -grad\, u$$

locally in Ω. If u is one-valued in Ω, for any pair of points $\boldsymbol{x}_1$,$\boldsymbol{x}_2 \in \Omega$ the e.m.f. V does not depend on the path $\Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2\ ) \subset \Omega$ but only on its endpoints and is equal to the voltage drop (or tension)

$$V = u(\boldsymbol{x}_1) - u(\boldsymbol{x}_2)$$

If in addition $\boldsymbol{E}$ is solenoidal,

$$\text{div}\, \boldsymbol{E} = 0$$

in Ω, the potential u is a harmonic function, i.e. satisfies the Laplace equation

$$div\, grad\, u \equiv \Delta_3 u = 0 \tag{L}$$

in Ω (see Exercise 6). The equipotential surfaces, defined in implicit form by $u =$ constant are obviously orthogonal to the lines of force of $\boldsymbol{E} = gradu$.

It is convenient to introduce, in addition to $\boldsymbol{E}$, a further field vector $\boldsymbol{D}$, called electric displacement vector, directly related to the electric charge by a universal law , i.e. a law valid in every instant of time and in every material medium, known as the Gauss Law. Denote by

$$Q[\Omega] = \sum_i Q_i + \int_\Omega \rho(\boldsymbol{x},t)\,\boldsymbol{dx} + \sum_j \int_{S_j} \sigma(\boldsymbol{x},t)\,dS_{\boldsymbol{x}} \tag{1.3}$$

the total charge contained in Ω, consisting of point charges Q_i, concentrated at points $\boldsymbol{x}_i \in \Omega$, surface charges distributed over surfaces $S_{\mathrm{j}} \subset \Omega$ with density $\sigma(\boldsymbol{x},t)$, and a volume charge distributed with density $\rho(\boldsymbol{x},t)$, $\boldsymbol{x}\in \Omega$ [2] Then the Gauss Law states that the total charge $Q[\Omega]$ contained inside an arbitrary (normal) domain $\Omega \subset \mathbb{R}^3$ is equal to the flux of D across $\partial\Omega$

$$\int_{\partial\Omega} \boldsymbol{D} \cdot \boldsymbol{n}\,dS = Q[\Omega] \tag{1.4}$$

This equation holds also if Ω is multiply connected.

The Gauss law implies that a point source Q at a point $\boldsymbol{x}_o$ in empty space generates a displacement vector field $\boldsymbol{D}$ that has a singularity at $\boldsymbol{x}_o$ and is described by Coulomb's law

$$\boldsymbol{D} = \frac{Q}{4\pi r^2}\frac{(\boldsymbol{x} - \boldsymbol{x}_o)}{r} \tag{1.5}$$

for $r = |\,\boldsymbol{x} - \boldsymbol{x}_{\,o}| > 0$. Coulomb's law is not universal and can be extended to dielectric media only under suitable assumptions. A dielectric or insulating material (e.g. air, wood, rubber, plastics...) is a material medium which does not conduct electricity and where charge distributions are permanent.

1.1.3 Conduction current.

In contrast to dielectrics, conductors are non-insulating media (e.g., metals) in which charges cannot be permanent but decay with time, giving rise to a conduction of electricity or flow of charge, called electric current.

[2] typically, the densities $\rho(\mathbf{x},t)$, $\sigma(\mathbf{x},t)$ are continuous and bounded functions

The electric current is described by the current density vector $\boldsymbol{J}$, defined in such a way that its lines of flows coincide with the lines of current, oriented according to the flow of positive charges, and the flux of $\boldsymbol{J}$ yields the current intensity I crossing any surface S

$$I = \int_S \boldsymbol{J} \cdot \boldsymbol{n}\, dS \tag{1.6}$$

It follows that, if $S = \partial\Omega$ is the boundary of a (normal) domain Ω, the current across $\partial\Omega$ is given by the integral over Ω of the divergence of $\boldsymbol{J}$

$$I = \int_\Omega div\, \boldsymbol{J}\, dV \tag{1.7}$$

A current tube (or tube of flow) is the set formed instantaneously by all the current lines (lines of flow of $\boldsymbol{J}$) that pass through a given surface S_o in the conductor. If

$$div\, \boldsymbol{J} \equiv 0$$

by applying (1.7) to a portion of tube of flow bounded by any two cross sections [3] S' and S'', we see that the tube of flow has "constant intensity", that is, the integral in (1.6) does not depend on the cross section S of the tube. Thus if $\boldsymbol{J}$ is solenoidal, the current $I = I(t)$ is constant along the tube for any fixed time t.

Lines of flow of a solenoidal vector field in a domain can be closed or begin and end at the boundary or at infinity, although more complicated geometries cannot be excluded. This applies to the geometry of lines of current whenever the density $\boldsymbol{J}$ is divergence-free.

The electric current I described by the density vector $\boldsymbol{J}$ is a spatially distributed or volume current, flowing in the current tube defined by the support of $\boldsymbol{J}$. When $div\, \boldsymbol{J} \equiv 0$ an important limiting case can be considered. Let $S = S(t)$ denote the cross section of maximum area $A = A(t)$ of the tube at any fixed time t. By the mean value theorem applied to the integral (1.6) we have

$$I = AJ \tag{1.8}$$

[3] a cross section is any open surface that disconnects the tube

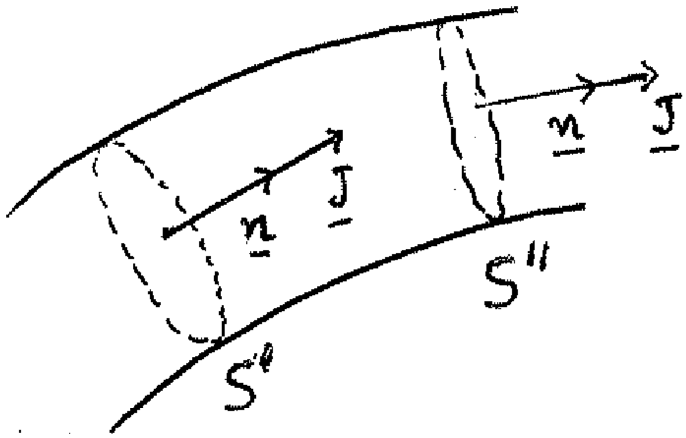

Figure 1.1: Current tube

where $J = J_S(t)$ is the value of $\boldsymbol{J}\cdot\boldsymbol{n}$ at some point of the cross section S at time t. Suppose that for any fixed time t the limit

$$I = \lim_{A\to 0}(AJ) \tag{1.9}$$

exists and is finite. We speak then of a linear current:the current tube shrinks in this limit to the central line of flow $\boldsymbol{x}=\boldsymbol{x}(s,t)$ modeling an infinitely thin conducting wire carrying the current $I = I(t)$, independent of the arclength s.

In linear currents, the density $\boldsymbol{J}$ is concentrated at the wire via a Dirac distribution, the so-called "delta function", and is directed along the tangent $\boldsymbol{t}$ to the wire. To clarify this point, consider a sequence S_n of shrinking cross sections, say with with areas $A_n = A/n\,(n = 1, 2, ...)$, and define the sequence $J_n = J_n(\boldsymbol{x}, t)$ by setting

$$J_n := \begin{cases} \frac{I}{A_n} & \text{if } \boldsymbol{x} \in S_n \\ 0 & \text{otherwise} \end{cases} \tag{1.10}$$

so that in the limit $A_n \to 0$ $J_n(\boldsymbol{x}, t)$ tends to zero for almost every $\boldsymbol{x}$ in S (i.e. excluding the point of intersection with the wire). Since $I = A_n J_n$ for any n, the limit (1.9) exists and is equal to $I = I(t)$; moreover the sequence $\boldsymbol{J}\cdot\mathbf{n} = J_n(\boldsymbol{x}, t)$ is such that

$$\lim_{n\to\infty} \int_S \boldsymbol{J} \cdot n \, dS = \lim_{n\to\infty} \int_{S_n} \frac{I}{A_n} dS = I$$

Thus, if $\boldsymbol{t}(s,t) = d\boldsymbol{x}/ds$ is the tangent to the wire at the point $\boldsymbol{x}(s,t)$, the weak limit $\overset{*}{\lim}_{n\to\infty}$ of the sequence $J_n(\boldsymbol{x}, t)$ identifies the Dirac distribution $\delta_\Gamma(\boldsymbol{x})$ concentrated at the wire Γ, and

$$\boldsymbol{J} = \overset{*}{\lim_{n\to\infty}} J_n \boldsymbol{t} = I\delta_\Gamma(\boldsymbol{x})\boldsymbol{t} \tag{1.11}$$

The physical consistency of this mathematical notion of an infinitely thin conducting wire will require a further discussion. In practice actual wires have a very small but finite cross section, and we define a linear current along such wires as the quantity I such that

$$\boldsymbol{J} \cong I\boldsymbol{t}/A$$

where A is the average cross section area and $\boldsymbol{t}$ is the tangent vector to the central line of flow of the wire. This definition is consistent since $div\, \boldsymbol{J} \equiv 0$.

Surface currents flowing over the boundary surface of "perfect conductors" are also possible, but we will never encounter situations where this notion is needed.

In a dielectric material there is no conduction of electric current, $\boldsymbol{J} \equiv \mathbf{0}$, and, as already mentioned, charges are permanent in time. From the physical (microscopic) point of view dielectrics and conductors are characterized by a different electron behavior: in dielectrics the electrons are bound to the ions, in conductors they are more or less free to move around and give rise to the conduction of electric current, with a mechanism which is similar, from a macroscopic point of view, to the phenomenon of heat conduction . The explanation of this mechanism, however, goes far beyond the aim and range of the macroscopic model provided by Maxwell's theory [4].

[4] of an entirely different nature is the convection current due to the motion of charged particles (Chapter 5)

1.1.4 Magnetic induction.

We pass now to the magnetic field, also described by two vector fields $\boldsymbol{B}$ and $\boldsymbol{H}$ satisfying universal laws. The magnetic field vector $\boldsymbol{B}$ (usually called magnetic induction) might in principle be defined, like the electric field $\boldsymbol{E}$, as the mechanical force acting over a unit positive magnetic pole. Such a definition, however, would be artificial, since it is well-known that a "magnetic charge", or magnetic pole, does not exist in nature [5]. On the other hand, magnetic dipoles do exist, in the shape, for example, of small bar magnets. The experiments show that a magnetic dipole under the influence of a magnetic field $\boldsymbol{B}$ in an arbitrary material medium undergoes a torque $\boldsymbol{T}$ given by

$$\boldsymbol{T} = \boldsymbol{m} \wedge \boldsymbol{B} \tag{1.12}$$

where the vector $\boldsymbol{m}$ describes a physical property of the dipole called (magnetic) moment (Exercise 7). The torque vanishes, and the dipole is in stable equilibrium, when $\boldsymbol{m}$ and $\boldsymbol{B}$ are parallel and with the same orientation: this explains why a compass needle tends to align itself to the Earth magnetic field (Exercise 8).

Similarly, a distribution $\boldsymbol{J}$ of electric current in a conductor in the presence of a magnetic field undergoes a force per unit volume $d\boldsymbol{F}/dV$ independent of the conductor and of the material medium surrounding it, and given by the "Ampère rule"

$$\frac{d\boldsymbol{F}}{dV} = \boldsymbol{J} \wedge \boldsymbol{B} \tag{1.13}$$

In particular the force acting over a rigid infinitely thin wire Γ carrying the linear current I (eq. (1.8)) is given by

$$\boldsymbol{F} = \int_{\mathbb{R}^3} \boldsymbol{J} \wedge \boldsymbol{B}\ dV = \int_{\mathbb{R}^3} I\boldsymbol{t} \wedge \boldsymbol{B}\ \delta_\Gamma(\boldsymbol{x})dV\ = I \int_\Gamma \boldsymbol{t} \wedge \boldsymbol{B}\ ds$$

where $\delta_\Gamma(\boldsymbol{x})dV$ is the Dirac measure concentrated at the wire. Thus

$$d\boldsymbol{F} = I\boldsymbol{t} \wedge \boldsymbol{B} ds \tag{1.14}$$

is the force due to $\boldsymbol{B}$ acting over a wire element ds . This equation implies an operative definition of $\boldsymbol{B}$ as the force divided by the current intensity times

[5] hence neither does a magnetic conduction current

the length of the element, and enables us to derive immediately the physical dimension unit of $\boldsymbol{B}$.

If the wire forms a very small (infinitesimal) rigid circular loop the vector $\boldsymbol{B}$ can be considered constant at all points of the loop, so that the integral in (1.14) becomes

$$\boldsymbol{F} = I \oint_\Gamma \boldsymbol{t} \wedge \boldsymbol{B}\, ds = -I\boldsymbol{B} \wedge \oint_\Gamma \boldsymbol{t}\, ds = -I\boldsymbol{B} \wedge \oint_\Gamma \frac{d\boldsymbol{x}}{ds} ds = \mathbf{0}$$

and the loop undergoes only a torque proportional to the current I and given by

$$\boldsymbol{T} = I\mathcal{A}\mathbf{b} \wedge \boldsymbol{B} \tag{1.15}$$

(Exercise 9), where $\mathcal{A} = \pi R^2$ is the (infinitesimal) area encompassed by the loop and $\mathbf{b}$ is the binormal vector, i.e. the normal to the plane of the loop oriented according to the right-handed screw rule with respect to $\boldsymbol{t}$ (Fig. 1.2). It follows that the circular loop tends to rotate around its diameter orthogonal to $\boldsymbol{B}$. This is essentially the principle underlying the operation of an electric motor.

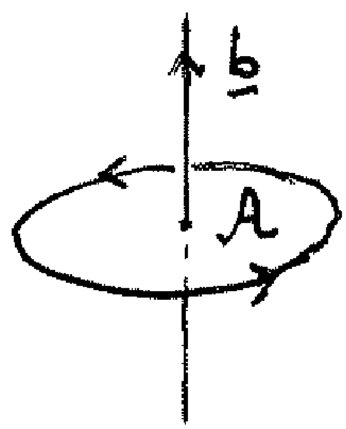

Figure 1.2: Right-handed screw rule

By comparing (1.15) with (1.12) we see that the loop behaves like a dipole with magnetic moment

$$\boldsymbol{m} = I\mathcal{A}\mathbf{b} \tag{1.16}$$

This equivalence led Ampère to envisage magnetism as electricity in motion.

The flux Φ of $\mathbf{B}$ across a generic surface S

$$\Phi := \int_S \boldsymbol{B} \cdot \boldsymbol{n}\, dS \tag{1.17}$$

will be seen to depend only on the boundary of S and is called the magnetic flux linking the circuit $\Gamma = \partial S$.

1.1.5 Magnetic field intensity.

As for the electric field, in order to formulate universal laws it is convenient to introduce a further magnetic vector $\boldsymbol{H}$, usually called magnetic field intensity. This vector field is directly generated by current-carrying wires: an element ds centered at a point $\boldsymbol{x}_o$ of an infinitely thin wire Γ carrying the current I generates at a point $\boldsymbol{x}$ of an arbitrary material medium a magnetic field intensity given by the Biot-Savart law

$$d\boldsymbol{H} = \frac{I}{4\pi} \frac{\boldsymbol{t} \wedge \boldsymbol{r}}{r^3}\, ds \tag{1.18}$$

where $\boldsymbol{r}{=}\boldsymbol{x}{-}\boldsymbol{x}_o$ and $\boldsymbol{t}$ is the tangent to the wire. This is the same law found in hydrodynamics for the fluid velocity distribution due to an element ds of a line vortex [8].

If $\boldsymbol{t}$ is constant along the wire, say $\boldsymbol{t}{=}\boldsymbol{t}_{\mathrm{o}}$, by integrating eq. (1.18) over s from $-\infty$ to $+\infty$ we obtain the Biot-Savart law for an unbounded rectilinear wire

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{I}{2\pi} \frac{\boldsymbol{t}_{\mathrm{o}} \wedge \boldsymbol{r}}{\varrho^2} \equiv \frac{I}{2\pi\varrho} \boldsymbol{\tau} \tag{1.19}$$

where $\varrho > 0$ is the distance of the point $\boldsymbol{x}$ from the wire, and $\boldsymbol{\tau} = \boldsymbol{t}_{\mathrm{o}} \wedge \boldsymbol{r}/|\boldsymbol{t}_{\mathrm{o}} \wedge \boldsymbol{r}|$ is the unit transverse vector in a plane orthogonal to $\boldsymbol{t}_{\mathrm{o}}$. If we choose $\boldsymbol{t}_o = \mathbf{c}_3$ and we introduce a polar coordinate system $\varrho = \sqrt{x_1^2 + x_2^2}$, $\varphi = \arctan(x_2/x_1)$ in the $(x_1, x_2)-$plane orthogonal to $\boldsymbol{t}_o$, we have

$$\boldsymbol{\tau} = -sin\varphi \mathbf{c}_1 + cos\varphi \mathbf{c}_2$$

and we can rewrite (1.19) in the form

$$H_1 = -\frac{I}{2\pi} \frac{x_2}{x_1^2 + x_2^2} \equiv -\frac{I\, sin\varphi}{2\pi\varrho} \quad , \quad H_2 = \frac{I}{2\pi} \frac{x_1}{x_1^2 + x_2^2} \equiv \frac{I\, cos\varphi}{2\pi\varrho} \quad , \quad H_3 = 0$$

so that

$$|\boldsymbol{H}| = \frac{I}{2\pi\varrho}$$

For $0 < a, b, \mathrm{M} < +\infty$ the integral

$$\int_{-a}^{b} dx_3 \int_0^{2\pi} d\varphi \int_0^{\mathrm{M}} |\boldsymbol{H}|^q \varrho d\varrho = \frac{I^{\,q}}{2^q \pi^q} \int_{-a}^{b} dx_3 \int_0^{2\pi} d\varphi \int_0^{\mathrm{M}} \frac{1}{\varrho^{q-1}} d\varrho$$

converges for $q = 1$ but diverges for $q = 2$. Therefore the singularity of $\boldsymbol{H}$ at the wire ($\varrho = 0$) is locally integrable in $\mathbb{R}^2$ and $\mathbb{R}^3$ but not locally square integrable.

Since the Biot-Savart field (1.19) is irrotational for $\varrho \neq 0$, there exists a magnetic potential, proportional to the polar angle φ

(1.20) $$v(\mathbf{x}) = -\frac{I}{2\pi}\varphi \equiv -\frac{I}{2\pi}\arctan{(x_2/x_1)}(r \neq 0)$$

such that $\boldsymbol{H} = -grad\, v$ for $\varrho \neq 0$. This potential v is many-valued, harmonic for $\varrho > 0$ and is not defined for $\varrho = 0$, in accordance with the fact that $\boldsymbol{H}(\boldsymbol{x})$ is irrotational for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \{z\text{−axis}\}$, a contourwise multiply connected domain. A branch surface for v is an arbitrary plane φ =constant and the period of v as a many-valued function [11] is equal to $-I$ (Exercise 10).

The magnetic intensity field generated by a magnetic dipole (a small magnetized bar) of moment $\boldsymbol{m}$ placed at a point $\mathbf{x}_o$ is given in an arbitrary material medium by

$$\boldsymbol{H}(\boldsymbol{x}) = -grad\, v(\boldsymbol{x})$$

where the potential v of the dipole is defined by

(1.21) $$v(\boldsymbol{x}) = \frac{\boldsymbol{m} \cdot \boldsymbol{r}}{4\pi r^3}$$

with $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}_o$, $r = |\boldsymbol{r}|$. Since v is harmonic for $r \neq 0$, the magnetic field $\boldsymbol{H}$ is irrotational and solenoidal for $r \neq 0$.

An electric dipole is described by the same formula (1.21) if $\boldsymbol{m}$ is interpreted as the electric dipole moment. The axis of a dipole is defined by the direction of $\mathbf{m}$.

The magnetomotive force (m.m.f.) is defined, in analogy to the e.m.f., by the circulation of $\boldsymbol{H}$

$$V_m = \int_\Gamma \boldsymbol{H} \cdot \boldsymbol{t}\, ds \equiv \sum_{k=1}^{3} \int_\Gamma H_k\, dx_k$$

and is independent of the path $\Gamma = \Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2)$ whenever $\boldsymbol{H}= -grad\, v$ with one-valued potential

Summarizing, we have seen that the pair of vectors $(\boldsymbol{E},\boldsymbol{B})$ is related to the forces or torques and the pair $(\boldsymbol{D},\boldsymbol{H})$ to the charges or currents, via universal laws valid in any material medium.

1.2 Maxwell equations in integral form

The Maxwell equations in their more straightforward form, namely the integral form, follow from four axioms that reflect universal phenomenological laws based on experiments and having unrestricted validity, irrespective of the material medium. We refer to begin with to bodies at rest . At rest with regard to what reference system? From the practical point of view we may take the laboratory system, where the observations were carried out and the experimental laws are valid; from a historical point of view, the privileged reference system was supposed to be at rest with respect to a hypothetical medium called (luminiferous) ether, pervading all space and conceived as the seat of electromagnetic phenomena. Unfortunately, the luminiferous ether does not exist. As we will discuss in Chap. 5, to formulate the electromagnetic laws for moving bodies by overcoming the concept of ether has been a fundamental stimulus towards the development of the restricted Relativity theory, which can be regarded as a consequence of Maxwell's equations [6] .

Let us consider a fixed closed curve Γ (independent of time) and an arbitrary fixed surface S having boundary $\partial S = \Gamma$. If $\boldsymbol{t}$ is the unit tangent vector along Γ, we fix the orientation of the normal $\boldsymbol{n}$ to S according to the right-handed screw rule.

The first axiom of Maxwell's theory reflects the Faraday induction law, which states that every time variation of the magnetic induction vector $\boldsymbol{B}$ gives rise to an induced electric field $\boldsymbol{E}$ such that the time rate of change of the magnetic flux Φ linking a closed path Γ is equal and opposite to the induced e.m.f. V in Γ :

$$\frac{d\Phi}{dt} = -V \tag{1.22}$$

[6] See F. Bampi, C.Zordan, Rend.Mat. 9, 417-425, 1989

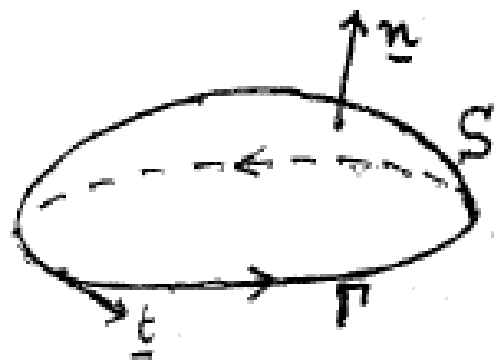

Figure 1.3: Orientation for the surface S

In other words we have, by (1.2) and (1.17),

$$-\frac{d}{dt}\int_S \boldsymbol{B}\cdot\boldsymbol{n}\; dS = \oint_\Gamma \boldsymbol{E}\cdot\boldsymbol{t}\,ds \qquad (\partial S = \Gamma). \tag{1.23}$$

where the left-hand side is independent of the choice of the surface S having Γ as border.

Viceversa, if the e.m.f. relative to a closed path Γ is different from zero, the magnetic flux Φ linking Γ necessarily varies in time. The induction law implies that Φ is stationary, i.e. $\frac{d\Phi}{dt}\equiv 0$, if $\boldsymbol{E}$ has a one-valued potential u, so that $V = 0$, or if the surface S is closed, so that $\partial S = \emptyset$.

In a conductor, the induced electric fields $\boldsymbol{E}$ gives rise to induced currents, called eddy currents in the case of bulk conductors.

Note that a time variation of Φ may also arise if $\boldsymbol{B}$ is stationary but S is in motion (Chapter 5). On the other hand, stationary magnetic fields do not induce electric currents in conducting bodies at rest.

If, instead of a single conducting loop, one considers a coil of N turns in series, eq. (1.22) becomes

$$V = -\sum_{k=1}^{N}\frac{d\Phi_k}{dt}$$

where Φ_k is the magnetic flux linking the kth turn. If they are all equal to

Φ, this becomes $V = -N d\Phi/dt,$ and the product $N\Phi$ is called the number of flux linkages.

The second axiom reflects the Ampère circuital law, which states that the m.m.f. V_m around any closed path Γ is equal to the total current linking the path :

$$V_m \ = I_{\text{tot}}$$

where

$$V_m \ := \oint_\Gamma \boldsymbol{H} \cdot \boldsymbol{t} \, ds$$

and the total current I_{tot}, defined as

$$I_{\text{tot}} \ = \ \int_S (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}) \cdot \boldsymbol{n} \ dS \qquad (\partial S = \Gamma)$$

is the sum of the conduction current $\boldsymbol{J}$ and the displacement current $\boldsymbol{J}_d$, given by

$$\boldsymbol{J}_d := \frac{\partial \boldsymbol{D}}{\partial t} \tag{1.24}$$

In other words, a conduction current $\boldsymbol{J}$ across a surface S and/or a time rate of change of $\boldsymbol{D}$ give rise to a magnetic intensity field $\boldsymbol{H}$ such that

$$\int_S (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}) \cdot \boldsymbol{n} \ dS = \oint_\Gamma \boldsymbol{H} \cdot \boldsymbol{t} \, ds \tag{1.25}$$

for every surface S such that $\partial S = \Gamma$. This implies that the total current vanishes if $\boldsymbol{H}$ has a one-valued potential v, so that $V_m = 0,$ or if S is a closed surface:

$$\oint_S (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}) \cdot \boldsymbol{n} \ dS = 0 \quad (\partial S = \emptyset) \tag{1.26}$$

Moreover, the total current through an open surface S depends only on ∂S and is called the total current "linking ∂S". The displacement current, one of the fundamental ideas of the Maxwell theory, is the sole current possible in a dielectric material, where $\mathbf{J} \equiv \mathbf{0}$.

The third axiom completes the first and states that, in accordance with the experience, the magnetic flux across any closed surface is zero

$$\oint_S \boldsymbol{B} \cdot \boldsymbol{n} \ dS = 0 \tag{1.27}$$

This equation implies that the magnetic flux through an open surface depends only on the surface boundary curve.

The fourth axiom is the Gauss law (1.4). By differentiating it with respect to time we obtain

$$\oint_{\partial\Omega} \frac{\partial \boldsymbol{D}}{\partial t} \cdot \boldsymbol{n}\, dS = \frac{dQ[\Omega]}{dt}$$

and taking eq. (1.26) into account with $S = \partial\Omega$ yields the equation of conservation of charge

$$\frac{dQ}{dt} + I = 0 \tag{1.28}$$

also called continuity equation. This balance equation states that for an arbitrary normal domain Ω the time rate of change of the total charge $Q = Q[\Omega]$ contained in the interior of Ω is equal to minus the outgoing current $I = I[\Omega]$ through the (conducting) boundary $\partial\Omega$. In this way the electric current is identified with the flux of electric charge. If Ω has an insulating boundary we have $I = 0$, hence $dQ/dt = 0$ and the total charge contained in Ω remains constant. The Faraday induction law and the Ampère circuital law can also be viewed as balance equations, for the e.m.f. V and the m.m.f. V_m, respectively,

$$V + \frac{d\Phi}{dt} = 0 \quad , \quad V_m - I_{\text{tot}} = 0$$

It is worth remarking that eq. (1.25) remains valid even if $\boldsymbol{H}$ has a many-valued potential v and an integrable singularity on S. For example, the lines of force of the stationary Biot-Savart magnetic field (1.19) for a rectilinear wire are circumferences Γ centered at the wire and the circulation along any one of them is

$$\oint_{\Gamma} \boldsymbol{H} \cdot \boldsymbol{t}\, ds = -\int_0^{2\pi} grad\, v \cdot \boldsymbol{t}\, \varrho d\varphi = \frac{I}{2\pi}\Big[\varphi\Big|_0^{2\pi} = I$$

where S is the circle bounded by Γ (Fig. 1.4).

Since I is given by

$$I \equiv \int_S \boldsymbol{J} \cdot \boldsymbol{n}\, dS$$

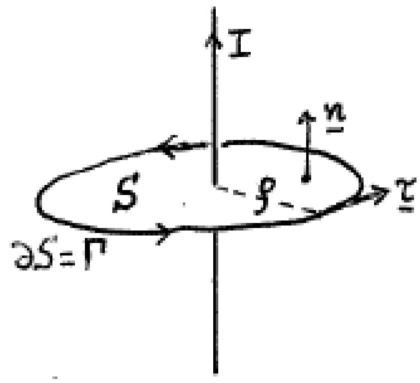

Figure 1.4: The Ampere circuital law for the Biot-Savart magnetic field

we see that (1.25) is satisfied (with $\frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{0}$) in spite of the fact that $\boldsymbol{H}$ has the integrable singularity $O(\varrho^{-1})$ at the center $\varrho = 0$ of S. On the other hand, the presence of this singularity follows from the fact that the circulation of $\boldsymbol{H}$ must be equal to $I \neq 0$ independently of the value of the radial distance ϱ. Thus the singularity of $\boldsymbol{H}$ at $\varrho = 0$ is intimately related to the many-valuedness of the potential v given by (1.20).

Similarly, eq. (1.23) remains valid if $\boldsymbol{E} = -grad\, u$ has a locally integrable singularity on S and the electric potential u is many-valued. Equations (1.3), (1.23), (1.25), (1.27) and (1.25), gathered together here

(I1)
$$\int_{\partial\Omega} \boldsymbol{D} \cdot \boldsymbol{n}\, dS = Q[\Omega]$$

(I2)
$$-\frac{d}{dt}\int_{S} \boldsymbol{B} \cdot \boldsymbol{n}\ dS = \int_{\partial S} \boldsymbol{E} \cdot \boldsymbol{t}\, ds$$

(I3)
$$\int_{S} (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}) \cdot \boldsymbol{n}\ dS = \int_{\partial S} \boldsymbol{H} \cdot \boldsymbol{t}\, ds$$

(I4)
$$\oint_{S} \boldsymbol{B} \cdot \boldsymbol{n}\ dS = 0$$

$$\frac{dQ}{dt} + I = 0 \tag{I5}$$

is the system of the Maxwell equations in integral form , valid for every normal domain Ω independent of time and for every surface S independent of time in $\mathbb{R}^3$. These equations are not all independent, since (I5) follows from (I3) and (I1) . Moreover, (I4) follows from (I2) if it holds at some fixed instant of time, say $t = 0$, and (I1) can be taken as a definition of the charge $Q = Q[\Omega]$. The first three equations (I1)−(I3) are independent but are clearly insufficient to determine all the unknowns.

In this integral formulation, directly related to experiments, there are no spatial derivatives and the vector fields ***E, D, B, H*** are allowed locally integrables singularities, in particular jump discontinuities across surfaces or curves in space.

1.3 Maxwell equations. Constitutive relations.

1.3.1 Maxwell field equations and constitutive relations.

Let us suppose that the charge Q is spatially distributed in a domain $\Omega \subset \mathbb{R}^3$ with continuous volume density $\rho(\boldsymbol{x}, t)$:

$$Q[\Omega] = \int_\Omega \rho(\boldsymbol{x}, t)\, \boldsymbol{dx} \tag{1.29}$$

so that $Q_i \equiv \sigma(x, t) \equiv 0$ in eq. (1.5). Moreover, suppose that ρ, ***E, D, B, H, J*** in eqs. (I1)−(I5) are of class $C^1(\mathbb{R}^4)$, that is, continuous functions of (x_1, x_2, x_3, t) together with all first partial derivatives in space−time. By applying Stokes's theorem (ST1) eqs. (I2), (I3) become

$$\int_S (curl\, \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t}) \cdot \boldsymbol{n}\, dS = 0 \quad , \quad \int_S (curl\, \boldsymbol{H} - \boldsymbol{J} - \frac{\partial \boldsymbol{D}}{\partial t}) \cdot \boldsymbol{n}\, dS = 0$$

and by the divergence theorem (DT) applied to eq. (I4) we obtain

$$\int_\Omega div\, \boldsymbol{B}\, dV = 0$$

Similarly, by applying the divergence theorem to eq. (I1) and by taking eq. (1.27) into account we find

$$\int_{\Omega} (div\, \boldsymbol{D} - \rho)\, dV = 0$$

Finally, by force of eqs. (1.6) and (1.29) the continuity equation (I5) takes the form

$$\int_{\Omega} (\frac{\partial \rho}{\partial t} + div\, \boldsymbol{J})\, dV = 0$$

All these integro-differential relations hold for any t, Ω and S :the mean value theorem, well known from Calculus, implies then that the integrands must vanish for all t and at all points of Ω and S. In this way we obtain the system of Maxwell's equations in differential form [7]

$$\text{(M1)} \qquad \frac{\partial \boldsymbol{B}}{\partial t} = -curl\, \boldsymbol{E}$$

$$\text{(M2)} \qquad \frac{\partial \boldsymbol{D}}{\partial t} = curl\, \boldsymbol{H} - \boldsymbol{J}$$

$$\text{(M3)} \qquad div \boldsymbol{D} = \rho$$

$$\text{(M4)} \qquad div \boldsymbol{B} = 0$$

$$\text{(M5)} \qquad \frac{\partial \rho}{\partial t} + div \boldsymbol{J} = 0$$

In this system the first three equations (M1)−(M3) are independent, eq. (M4) has the character of an initial condition, and the continuity equation (M5) follows from (M2) and (M3). Indeed, equation (M1) implies that

$$\boldsymbol{B}(\boldsymbol{x}, t) = \boldsymbol{B}(\boldsymbol{x}, t_o) - curl \int_{t_o}^{t} \boldsymbol{E}(\boldsymbol{x}, \tau)\, d\tau \tag{1.30}$$

[7] This form is due to Heaviside, see B.J. Hunt "The Maxwellians", Cornell University Press 1991.

whence

$$div\boldsymbol{B}(\boldsymbol{x},t) = div\boldsymbol{B}(\boldsymbol{x},t_o)$$

so that eq. (M4) holds for all t if it holds at some fixed (say, initial) time t_o. Similarly, the continuity equation (M5) follows by taking the divergence of (M2) and by applying (M3). In its turn, eq. (M3) can be used to eliminate the unknown ρ, by defining the electric volume charge density in terms of $\boldsymbol{D}$ as

$$\rho := div\,\boldsymbol{D}$$

In this way, the Maxwell system reduces to the two vector equations (M1), (M2) , valid in any material medium, conducting or non-conducting, for the five unknown vector functions $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{B}$, $\boldsymbol{H}$, $\boldsymbol{J}$ of $(\boldsymbol{x},t)$. These two vector equations are complemented by three additional vector relations, called constitutive equations, and so the count is right. These constitutive relations are not universally valid but depend upon the properties of the materials under consideration. We can assume to start with that they have the form of local relations

$$\boldsymbol{J} = \boldsymbol{J}(\boldsymbol{E},\boldsymbol{H})$$
$$\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{E},\boldsymbol{H})$$
$$\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{E},\boldsymbol{H})$$

and in fact for many purposes we will take the very simple linear constitutive relations

(C1) $$\boldsymbol{J} = \gamma\,\boldsymbol{E} \qquad \text{(Ohm's law)}$$

(C2) $$\boldsymbol{D} = \epsilon\,\boldsymbol{E}$$

(C3) $$\boldsymbol{B} = \mu\,\boldsymbol{H}$$

where $\gamma = \gamma(\boldsymbol{x}) \geq 0$ is the electric conductivity, γ^{-1} the resistivity, $\epsilon = \epsilon(\boldsymbol{x}) \geq \epsilon_o > 0$ is the electric permittivity and $\mu = \mu(\boldsymbol{x}) \geq 0$ the magnetic permeability of the material. These relations apply to empty space with $\epsilon = \epsilon_o$, $\mu = \mu_o$, $\gamma = 0$ and the more common materials can be classified according to the values of the scalar coefficients ϵ, μ, γ as follows:

(1.31) $$\begin{cases} \gamma = 0 & : \text{ dielectrics} \\ 0 < \gamma < \infty & : \text{ conductors} \\ \gamma = +\infty & : \text{ perfect conductors} \end{cases}$$

$$\begin{cases} \mu > \mu_o & : \quad \text{paramagnetic bodies} \\ 0 < \mu < \mu_o & : \quad \text{diamagnetic bodies} \\ \mu = 0 & : \quad \text{superconductors} \end{cases} \tag{1.32}$$

where ϵ_o, μ_o are the (constant) permittivity and permeability of empty space. We will exclude in the sequel the case of superconductors [8] and we will always assume that there exists $\overline{\mu} > 0$ such that

$$\mu(\boldsymbol{x}) \geq \overline{\mu} > 0 \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^3$$

For homogeneous media the coefficients γ, ϵ and μ are constant. They depend on physical parameters such as temperature: for example, the conductivity of metals decreases with increasing temperature [22].

The linear constitutive relations (C1)$-$(C3) apply to a wide class of isotropic materials if the field intensities are not too high and the temperature is not too low; in particular, Ohm's law (C1) holds in the static case and for not too high frequencies (up to the infrared for typical metals at room temperatures). In contrast, in the case of anisotropic bodies such as crystals the permittivity becomes a 3×3 rank-2 tensor ϵ_{jk}, and for very intense electric fields nonlinear effects in the dependence of $\boldsymbol{D}$ upon $\boldsymbol{E}$ become important, giving rise to phenomena pertaining to the field of nonlinear optics.

For ferromagnetic bodies, such as iron or steel, the constitutive relation $\boldsymbol{B}=\boldsymbol{B(H)}$ is nonlinear and possibly many-valued (due to hysteresis), and the magnetic permeability can be defined in an extended sense only in the case of the so-called "magnetically soft bodies" and is a function of $|\boldsymbol{H}|$. On the other hand, for non-ferromagnetic bodies $\mu(\boldsymbol{x})$ is practically constant and can be taken equal to μ_o. In this case we speak of non-magnetic bodies. Such materials are characterized by the fact that a stationary magnetic field $\boldsymbol{B}$ exerts no forces on them if they are at rest and carry no current.

Strictly speaking, from a physical point of view $\boldsymbol{D}$ depends also on past values of $\boldsymbol{E}$ via a nonlocal "hereditary" relation of convolution type which in the case of isotropic bodies has the form

$$\boldsymbol{D}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\epsilon}(\boldsymbol{x}, \tau) \boldsymbol{E}(\boldsymbol{x},\, t - \tau) d\tau \tag{HR}$$

[8] a discussion on superconductors can be found in [35], ch. 6

with $\tilde{\epsilon}(\boldsymbol{x},\tau) \equiv 0$ for $\tau < 0$ (see e.g. [35], p. 388 sgg.). For dispersive anisotropic bodies the memory function $\tilde{\epsilon}(\boldsymbol{x},\tau)$ is replaced by a memory tensor $\tilde{\epsilon}_{ij}(\mathrm{x},\tau)$. For periodic phenomena, the convolution theorem for Fourier integrals [3] implies that the permittivity for each Fourier component depends on the frequency. A similar remark can be made concerning Ohm's law (C1) (See Chapter 6.) Thus the linear local constitutive relations (C1),(C2) must be viewed as an approximate model of physical reality.

To summarize, the Maxwell equations consist of two well distinct groups of equations, the linear field equations (M1)−(M4), having universal validity, and the constitutive relations, which depend on the particular material media under consideration, and in certain cases may be nonlinear and/or nonlocal. By adding the linear constitutive relations (C1), (C2), (C3) to the equations (M1) and (M2) we obtain the closed set of five vector equations in the five vector unknown ***E, D, B, H, J***

$$\frac{\partial \boldsymbol{B}}{\partial t} = -curl\,\boldsymbol{E}\,,\ \frac{\partial \boldsymbol{D}}{\partial t} = curl\,\boldsymbol{H} - \boldsymbol{J},\quad \boldsymbol{D} = \epsilon\boldsymbol{E},\quad \boldsymbol{J} = \gamma\boldsymbol{E},\ \boldsymbol{B} = \mu\boldsymbol{H} \tag{1.33}$$

with $div\,\boldsymbol{B} = 0$ (initially) and volume charge density defined by $\rho = div\boldsymbol{D}$. The coefficients γ, ϵ and μ will be assumed to be C^1 functions as the point $\boldsymbol{x}$ varies inside any single material medium, with possible jump discontinuities at the boundaries between different media.

Since ***D, B*** and $\boldsymbol{J}$ can always be expressed in terms of $\boldsymbol{E}$ and $\boldsymbol{H}$ using the constitutive relations, it is customary to think of $\boldsymbol{E}$, $\boldsymbol{H}$ as the fundamental electromagnetic field and of $\boldsymbol{D}$,$\boldsymbol{B}$,$\boldsymbol{J}$ as derived quantities.

Unless stated otherwise, in Chapters 1−5 we will always consider isotropic materials for which the linear constitutive relations (C1)−(C3) are valid. The electromagnetic field ($\boldsymbol{E}$,$\boldsymbol{H}$) satisfies then the system

$$\mu\frac{\partial \boldsymbol{H}}{\partial t} = -curl\boldsymbol{E} \qquad , \qquad \epsilon\frac{\partial \mathbf{E}}{\partial t} = curl\boldsymbol{H} - \gamma\boldsymbol{E}$$

obtained from eqs.(1.33) by eliminating ***D, B, J***. Moreover, we will often assume that the materials are homogeneous, so that ϵ, μ and γ are constant.

1.3.2 Some consequences of the linear constitutive relations.

Let us now examine a few consequences of the constitutive relations (C1)-(C3).

We begin by discussing Ohm's law (C1). It implies that the lines of flow of electric current density **J** in a conductor coincide with the lines of electric field intensity **E** and are perpendicular to equipotential surfaces, if an electric potential exists. This is by no means an evident fact. Consider a straight homogeneous cylindrical conducting wire of length l with uniform cross section S of very small but finite area A, carrying a steady current I, parallel to a constant unit vector $\boldsymbol{t}_o$. Each cross section of the wire has then constant normal vector $\boldsymbol{n}{=}\boldsymbol{t}_o$ parallel to the lines of current. By Ohm's law (C1), $\boldsymbol{J}= \gamma \boldsymbol{E}$ where $\boldsymbol{J}$ is uniform and is given by

$$\boldsymbol{J} = \frac{I\boldsymbol{t}}{A} \tag{1.34}$$

with $\boldsymbol{t}{=}\boldsymbol{t}_o$. Therefore the electric field $\boldsymbol{E}$ inside the wire is also uniform and directed along the lines of current.Integrating $\boldsymbol{J}\cdot\boldsymbol{n}= I/A$ and $\gamma\boldsymbol{E}\cdot\boldsymbol{n}= \gamma\boldsymbol{E}\cdot\boldsymbol{t}_o$ over the cylinder $(0, l) \times S$ we find

$$\int_0^l \int_S \boldsymbol{J}\cdot\boldsymbol{n}\, dSds = Il, \quad \gamma \int_0^l \int_S \boldsymbol{E}\cdot\boldsymbol{n}\, dSds = \gamma A \int_0^l \boldsymbol{E}\cdot\boldsymbol{t}_o\, ds = \gamma AV$$

where V is the e.m.f. By Ohm's law the two integrals must coincide, and we find Ohm's law for linear currents

$$V = I\mathcal{R}$$

where

$$\mathcal{R} = \frac{l}{\gamma A} \tag{1.35}$$

is the wire resistance, proportional to its length l. $\mathcal{R}$ is also inversely proportional to the cross-section area A and diverges as A tends to zero, so that the limit of an infinitely thin wire has no meaning in this context.

The Faraday induction law (1.22) combined with Ohm's law $V = I\mathcal{R}$ in the case of linear currents takes then the form

$$\frac{d\Phi}{dt} = -\mathcal{R}I \tag{1.36}$$

where I is the linear current induced in the circuit by the rate of change of magnetic flux $d\Phi/dt$. More generally, the constitutive relation $\boldsymbol{J}= \gamma \boldsymbol{E}$ combined with the Faraday induction law (1.22) implies that every time variation of the magnetic flux Φ in a conductor gives rise to an induced current $\boldsymbol{J}$ given by $\boldsymbol{J}= \gamma \boldsymbol{E}$, where $\boldsymbol{E}$ is the induced electric field. The induced current (called eddy current in the case of bulk conductors) is always established in such a direction that its magnetic field, produced according to the Ampère circuital law, tends to oppose the change of Φ. This experimental result, called Lenz's law, agrees with the Faraday induction law (the verification is left as an exercise). Lenz's law was confirmed experimentally in 1824 by Arago and Faraday who observed that the rotation of a horizontal copper disk provoked a companion rotation in a magnet (a small compass needle) mounted coaxially in the proximity of the disk. Indeed, the eddy currents arising in the copper disk tend to oppose the change in magnetic flux, thus driving the needle to rotate with the same angular speed (Fig. 1.5).

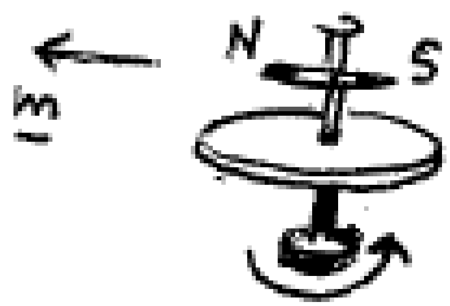

Figure 1.5: Experimental confirmation of Lenz's law

We consider next the constitutive equation (C2), $\boldsymbol{D}= \epsilon \boldsymbol{E}$. In empty space, $\boldsymbol{D}= \epsilon_o \boldsymbol{E}$ and Coulomb's law (1.5) combined with eq. (1.1) implies that a point source Q at a point $\boldsymbol{x}_o$ generates a field of force at a point $\mathbf{x}$ in empty space given by $\boldsymbol{E}(\boldsymbol{x}) = \epsilon_o^{-1}\boldsymbol{D}(\boldsymbol{x})$, i.e.

$$\boldsymbol{E} = \frac{Q}{4\pi\epsilon_o r^2}\,\frac{(\boldsymbol{x}-\boldsymbol{x}_o)}{|\boldsymbol{x}-\boldsymbol{x}_o|}$$

and $q\boldsymbol{E}$ is then (according to (1.1)) the Coulomb force acting over a test charge q concentrated at $\boldsymbol{x}$, due to the source charge Q concentrated at $\boldsymbol{x}_o$. Note that the Coulomb force exerted by Q over Q itself vanishes due to symmetry, so that there is no self-induced action. This fact can be seen by computing the self-induced force over a uniform volume distribution of charge Q in a ball B_R of radius R centered at $\mathbf{x}_o$, and then passing to the limit as R goes to zero. By applying Proposition 1.3.4 below and eq. (C1) with $\epsilon = \epsilon_o$ we have for $|\boldsymbol{x}-\boldsymbol{x}_o| < R$

$$\boldsymbol{E}(\boldsymbol{x}) = \frac{\rho_o(\boldsymbol{x} - \boldsymbol{x}_o)}{3\epsilon_o} \quad , \qquad \rho_o = \frac{Q}{\frac{4}{3}\pi R^3}$$

and the corresponding self-induced force, given by the integral

$$\boldsymbol{F}_R = \int_{B_R} \rho_o \boldsymbol{E} d\boldsymbol{x}$$

vanishes for all R due to symmetry, so that the limit of $\boldsymbol{F}_R$ as $R \to 0$ is zero too.

Similar considerations can be carried out with regard to the magnetic field

$$\boldsymbol{H} = -grad v = \frac{1}{4\pi r^5} \sum_{i=1}^{3} r_i(\boldsymbol{m} r_i - 3m_i \boldsymbol{r}) \qquad \boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}_o \tag{1.37}$$

due to a magnetic dipole concentrated at a point $\boldsymbol{x}_o$ in a homogeneous medium with permeability μ. According to eq. (1.12), the corresponding magnetic induction field $\boldsymbol{B}= \mu\boldsymbol{H}$ might in principle exert a self-induced torque $\boldsymbol{T}=\boldsymbol{m}\wedge\boldsymbol{B}$ over the dipole itself. A limit argument again shows that this self-induced torque is zero for reasons of symmetry.

The issue of self-induced action in the case of the Biot-Savart magnetic field due to an infinitely thin electric wire is more involved. In the case of two non-intersecting wires Γ_j carrying the currents I_j ($j = 1, 2$), the Biot-Savart law (1.18) shows that the total magnetic field due to two elements ds_1, ds_2 centered at the points $\boldsymbol{x}_1 \in \Gamma_1$ and $\boldsymbol{x}_2 \in \Gamma_2$, respectively, is given by $d\boldsymbol{H}_1 + d\boldsymbol{H}_2$, with

$$d\boldsymbol{H}_j(\boldsymbol{x}) = \frac{I_j}{4\pi} \frac{\boldsymbol{t}_j \wedge (\boldsymbol{x} - \boldsymbol{x}_j)}{|\boldsymbol{x} - \boldsymbol{x}_j|^3} ds_j \qquad (j = 1, 2)$$

The force acting over the first wire element due to the magnetic field $d\boldsymbol{B}_2 = \mu d\,\boldsymbol{H}_2$ of the second is given by eq. (1.14)

$$d\boldsymbol{F}_{12} = I_1 \boldsymbol{t}_1 \wedge d\boldsymbol{B}_2 ds_1 = \frac{I_1 I_2}{4\pi} \frac{\boldsymbol{t}_1 \wedge (\boldsymbol{t}_2 \wedge (\boldsymbol{x}_1 - \boldsymbol{x}_2))}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3} ds_1 ds_2$$

and the force $d\boldsymbol{F}_{21}$ exerted over the second wire element by the magnetic field $d\boldsymbol{B}_1 = \mu d\boldsymbol{H}_1$ of the first is

$$d\boldsymbol{F}_{21} = I_2 \mathbf{t}_2 \wedge d\boldsymbol{B}_1 ds_2 = \frac{I_2 I_1}{4\pi} \frac{\boldsymbol{t}_2 \wedge (\boldsymbol{t}_1 \wedge (\boldsymbol{x}_2 - \boldsymbol{x}_1))}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3} ds_2 ds_1$$

If $\boldsymbol{t}_1 = \boldsymbol{t}_2$, i.e. if the two wires are parallel, $d\boldsymbol{F}_{12} = -d\boldsymbol{F}_{21}$ and this mutual action is attractive when $I_1 I_2 > 0$ and repulsive otherwise. This fact was confirmed experimentally by Ampère.

Oersted discovered (1820) that an electric current exerts a mechanical torque $\mathbf{T}$ over a magnet by mounting a compass needle free to rotate around a vertical axis just beneath a straight horizontal conducting wire. If the needle is parallel to the wire then a torque proportional to the current I arises which causes the needle to rotate around its axis; the torque vanishes when the needle is orthogonal to the wire. This phenomenon is easily explained by modelling the needle as a magnetic dipole with moment $\boldsymbol{m}$ and by applying eq. (1.12) for the torque over a magnet, together with the Biot-Savart law (1.19) and the constitutive relation (C3). The result is

$$\boldsymbol{T} = \boldsymbol{m} \wedge \mu \boldsymbol{H} = \frac{\mu\, I}{2\pi d}\ \boldsymbol{m} \wedge \boldsymbol{\tau}$$

where d is the distance between the wire and the needle and $\boldsymbol{\tau}$ is a horizontal unit vector orthogonal to the wire. Thus the torque vanishes only when $\boldsymbol{m}$ is aligned $\boldsymbol{H}$, i.e. when the needle is orthogonal to the wire.

Further consequences of the linear, local constitutive relations can be stated as corollaries to the Maxwell equations.

Proposition 1.3.1 *In a homogeneous conductor all charges decay in time, permanent (steady) charges are impossible, and the relation*

$$\gamma \rho(\boldsymbol{x}) \equiv 0$$

holds in steady conditions.

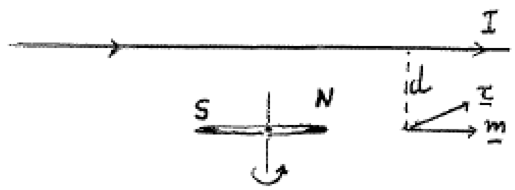

Figure 1.6: Oersted's experiment

Proof. If $\gamma > 0$ and the permittivity ϵ is a scalar constant, substituting the constitutive relations (C1), (C2) into eq. (M5) yields

$$0 = \frac{\partial\rho}{\partial t} + \ div\boldsymbol{J} \ = \frac{\partial\rho}{\partial t} + \gamma\, div\boldsymbol{E} = \frac{\partial\rho}{\partial t} + \frac{\gamma}{\epsilon}\, div\boldsymbol{D}$$

and by virtue of (M3) we find the differential equation for ρ

$$\frac{\partial\rho}{\partial t} + \frac{\gamma}{\epsilon}\rho \ = 0$$

Hence $\partial\rho/\partial t \equiv 0$ implies $\gamma\rho \equiv 0$, as asserted. By solving this differential equation we obtain

$$\rho = \rho_o(\boldsymbol{x})\, exp\,(-\frac{\gamma}{\epsilon}t) \equiv \rho_o(\boldsymbol{x})\, exp\,(-t/\tau)$$

where $\tau = \epsilon/\gamma$ is the relaxation time of the conductor, and $\rho_o(\boldsymbol{x})$ is the initial density. Similarly div $\boldsymbol{J} = \tau^{-1}\rho_o(\boldsymbol{x})e^{-t/\tau}$, $Q(t) = Q_o e^{-t/\tau}$.

In dielectrics $\gamma = 0$, $\tau = +\infty$, and charges are permanent. In contrast, for perfect conductors $\gamma = +\infty$, $\tau = 0$, and we have

Proposition 1.3.2 *Inside a perfect conductor* $\boldsymbol{E} \equiv \boldsymbol{D} \equiv \rho \equiv 0$ *and* $\boldsymbol{B}$ *is stationary.*

Proof. $\gamma = +\infty$ implies $\boldsymbol{E} \equiv 0$ because of (C1), $\boldsymbol{D} \equiv 0$ by (C2), $\rho \equiv 0$ by (M3), and $\partial\boldsymbol{B}/\partial t \equiv \boldsymbol{0}$ by (M1).

Proposition 1.3.3 *Maxwell's equations imply Coulomb's law extended to an arbitrary homogeneous unbounded dielectric medium with permittivity* ϵ :

$$\boldsymbol{E}(\boldsymbol{x}) = \frac{Q}{4\pi\epsilon r^2}\ \frac{\boldsymbol{x} - \boldsymbol{x}_o}{r}\ ; \quad \boldsymbol{E} = -grad\, u\ , \quad u(\boldsymbol{x}) = \frac{Q}{4\pi\epsilon r} \tag{1.38}$$

Proof. Consider a single point charge Q concentrated at a point $\boldsymbol{x}_o$ of an unbounded dielectric material medium with constant permittivity ϵ occupying all of $\mathbb{R}^3$. For symmetry reasons $\boldsymbol{D}(\boldsymbol{x}) = D(r)\boldsymbol{n}$, $\boldsymbol{n}= (\boldsymbol{x}-\boldsymbol{x}_o)/|\boldsymbol{x}-\boldsymbol{x}_o|$. The Gauss law (I1) applied to the sphere $S : |\boldsymbol{x}-\boldsymbol{x}_o| = r$ yields then

$$Q = \int_S \mathbf{D} \cdot \boldsymbol{n}\, dS = D(r) 4\pi r^2$$

Hence for every $r > 0$

$$\boldsymbol{D}(\boldsymbol{x}) = \frac{Q}{4\pi r^2} \frac{\boldsymbol{x} - \boldsymbol{x}_o}{|\boldsymbol{x} - \boldsymbol{x}_o|} \tag{1.39}$$

and $\boldsymbol{E}= \epsilon^{-1}\boldsymbol{D}$ is given by eq. (1.38), is irrotational, and its potential is $u = \frac{Q}{4\pi\epsilon r}$+constant.

Proposition 1.3.4 *A uniform volume charge distribution* ρ_o *in a ball* B_R *of radius* R *centered at* $\boldsymbol{x}_o$ *and immersed in a homogeneous unbounded dielectric gives rise to a displacement field* $\boldsymbol{D}$ *given by*

$$\boldsymbol{D}(\boldsymbol{x}) = \frac{\rho_o(\boldsymbol{x} - \boldsymbol{x}_o)}{3} \qquad \text{for } r = |\boldsymbol{x} - \boldsymbol{x}_o| \leq R$$

and by eq. (1.39) with $Q = \frac{4}{3}\pi R^3 \rho_o$ *for* $r = |\boldsymbol{x}-\boldsymbol{x}_o| > R.$

Proof. For symmetry reasons $\boldsymbol{D}(\boldsymbol{x}) = D(r)\boldsymbol{n}$ as in Proposition 1.3.3, and the Gauss law applied to the sphere S_r of radius r yields

$$D(r) = \frac{1}{4\pi r^2} \int_{S_r} \boldsymbol{D} \cdot \boldsymbol{n}\, dS = \begin{cases} Q/4\pi r^2 & \text{if} \quad \mathrm{r} > R \\ \frac{4}{3}\pi r^3 \rho_o / 4\pi r^2 = \rho_o r/3 & \text{if} \quad r < R \end{cases}$$

Since $r\boldsymbol{n}=\boldsymbol{x}-\boldsymbol{x}_o$ and $Q = \frac{4}{3}\pi R^3 \rho_o$, $\boldsymbol{D}$ behaves as stated and is continuous for $r = R$.

Proposition 1.3.5 *Similarly, if the charge* Q *is distributed over a sphere of radius* R *centered at* $\boldsymbol{x}_o$ *with constant surface density* $\sigma_o = Q/4\pi R^2$, *then* $\boldsymbol{D}$ *is zero in the interior of the sphere and is given by eq. (1.39) on the exterior.*

Proof. The total charge contained in a sphere of radius r is $Q = 4\pi R^2\sigma_o$ if $r > R$, 0 if $r < R$. The same argument as in Proposition 1.3.4 then shows that

$$\boldsymbol{D}(r) = \frac{1}{4\pi r^2}\int_{S_r} \boldsymbol{D}\cdot\boldsymbol{n}\, dS = \begin{cases}\sigma_o & \text{if} \quad \mathrm{r} > R\\ 0 & \text{if} \quad r < R\end{cases}$$

and $\boldsymbol{D}= D(r)\,(\boldsymbol{x}-\boldsymbol{x}_o)/|\boldsymbol{x}-\boldsymbol{x}_o|$ is now discontinuous on the surface of the sphere $r = R$ if $\sigma_o \neq 0$.

The Biot-Savart law (1.18) also follows from Maxwell's equations, as will be seen later on.

1.4 Matching conditions

Consider two different media occupying the domains Ω_-, Ω_+ separated by a smooth surface $\mathbb{S}$ independent of time. Suppose that finite limits γ_-, ϵ_- , μ_- and γ_+, ϵ_+ , μ_+ exist for $\gamma(\boldsymbol{x}), \epsilon(\boldsymbol{x})$ and $\mu(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches $\mathbb{S}$ in Ω_- and Ω_+, respectively. In general these limits will be different so that the Maxwell equations (1.33) will have discontinuous coefficients $\gamma(\boldsymbol{x})$, $\epsilon(\boldsymbol{x}), \mu(\boldsymbol{x})$ in $\mathbb{R}^3$ with jump discontinuities across $\mathbb{S}$. The solution vectors ***E, H, B, D, J*** may then well be discontinuous on $\mathbb{S}$. If this happens, the differential form of the Maxwell equations holds only where the fields are smooth, i.e. in the interior of Ω_- and Ω_+, whereas the integral form remains valid in all of $\mathbb{R}^3$ and yields informations about the behavior of ***E, H, B, D, J*** across the surface $\mathbb{S}$.

Suppose that $\mathbb{S}$ is smooth (say, of class C^2) and that the normal $\boldsymbol{n}$ is oriented from Ω_- to Ω_+ . Moreover, suppose that the vector fields $\boldsymbol{E}(\boldsymbol{x}, t)$, $\boldsymbol{H}(\boldsymbol{x}, t)$, $\boldsymbol{J}(\boldsymbol{x}, t)$,$\boldsymbol{B}(\boldsymbol{x}, t)$, $\boldsymbol{D}(\boldsymbol{x}, t)$ are bounded together with their first derivatives and that the limits as $\mathbf{x}$ approaches $\mathbb{S}$ in Ω_+ and Ω_-

$$\boldsymbol{E}_\pm(\boldsymbol{y}, t) := \lim_{h\to 0^+} \boldsymbol{E}(\boldsymbol{y} \pm h\boldsymbol{n}(\boldsymbol{y}), t)$$

exist for all $\boldsymbol{y}\in \mathbb{S}$ and are continuous on $\mathbb{S}$ for every t (similarly for $\boldsymbol{H}$, $\boldsymbol{J}$,$\boldsymbol{B}$, $\boldsymbol{D}$). Let

$$[\boldsymbol{E}]_{\mathbb{S}} := \boldsymbol{E}_+ - \boldsymbol{E}_-$$

denote the jump at the point $\mathbf{y}\in \mathbb{S}$, defined as the difference of the two limits as $\boldsymbol{x}$ approaches $\boldsymbol{y}\in \mathbb{S}$. By assumption, $[\boldsymbol{E}]_{\mathbb{S}}$ is continuous on $\mathbb{S}$ for every

t. Let $\boldsymbol{y} \in \mathbb{S}$ be an arbitrary point of $\mathbb{S}$ and $\boldsymbol{n}=\boldsymbol{n}(\boldsymbol{y})$, $\sigma = \sigma(\boldsymbol{y}, t)$ the unit normal and the surface charge density at the point $\boldsymbol{y} \in \mathbb{S}$, respectively. Then the following matching relations hold for every t :

(R1) $$[\boldsymbol{D}]_{\mathbb{S}} \cdot \boldsymbol{n} = \sigma$$

(R2) $$[\boldsymbol{B}]_{\mathbb{S}} \cdot \boldsymbol{n} = 0$$

(R3) $$[\boldsymbol{J}]_{\mathbb{S}} \cdot \boldsymbol{n} + \frac{\partial \sigma}{\partial t} = 0$$

(R4) $$[\boldsymbol{H}]_{\mathbb{S}} \wedge \boldsymbol{n} = \boldsymbol{0}$$

(R5) $$[\boldsymbol{E}]_{\mathbb{S}} \wedge \boldsymbol{n} = \boldsymbol{0}$$

In other words, the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$

(1.40) $$\boldsymbol{E}_{tang} = -(\boldsymbol{E} \wedge \boldsymbol{n}) \wedge \boldsymbol{n} \quad , \quad \boldsymbol{H}_{tang} = -(\boldsymbol{H} \wedge \boldsymbol{n}) \wedge \boldsymbol{n}$$

and the normal component $\boldsymbol{B} \cdot \boldsymbol{n}$ of $\boldsymbol{B}$ are continuous across $\mathbb{S}$. In contrast, the normal components $\boldsymbol{D} \cdot \boldsymbol{n}$, $\boldsymbol{J} \cdot \boldsymbol{n}$ have a discontinuity jump equal to σ and $-\partial\sigma/\partial t$, respectively: this is to be interpreted in the sense that, for any fixed $\boldsymbol{y} \in \mathbb{S}$ and any t,

$$\lim_{h \to 0^+} \left[\boldsymbol{D}(\boldsymbol{y} + h\boldsymbol{n}(\boldsymbol{y}), t) \cdot \boldsymbol{n}(\boldsymbol{y}) - \boldsymbol{D}(\boldsymbol{y} - h\boldsymbol{n}(\boldsymbol{y}), t) \cdot \boldsymbol{n}(\boldsymbol{y})\right] = \sigma(\boldsymbol{y}, t)$$

$$\lim_{h \to 0^+} \left[\boldsymbol{J}(\boldsymbol{y} + h\boldsymbol{n}(\boldsymbol{y}), t) \cdot \boldsymbol{n}(\boldsymbol{y}) - \boldsymbol{J}(\boldsymbol{y} - h\boldsymbol{n}(\boldsymbol{y}), t) \cdot \boldsymbol{n}(\boldsymbol{y})\right] = -\frac{\partial \sigma(\boldsymbol{y}, t)}{\partial t}$$

In order to prove (R1), let $\boldsymbol{x}_o$ be an arbitrary point of $\mathbb{S}$, $\mathbb{I}_o$ denote a circular neighborhood of $\boldsymbol{x}_o$ with radius ε $(0 < \varepsilon \ll 1)$ on $\mathbb{S}$, and

$$\Omega_h = (-h, h) \times \mathbb{I}_o \qquad , \qquad 0 < h \ll 1$$

be the cylindrical pillbox having generatrices parallel to $\boldsymbol{n}_o = \boldsymbol{n}(\boldsymbol{x}_o)$ and the two bases in Ω_- and Ω_+ on opposite sides of $\mathbb{S}$ (Fig. 1.7).

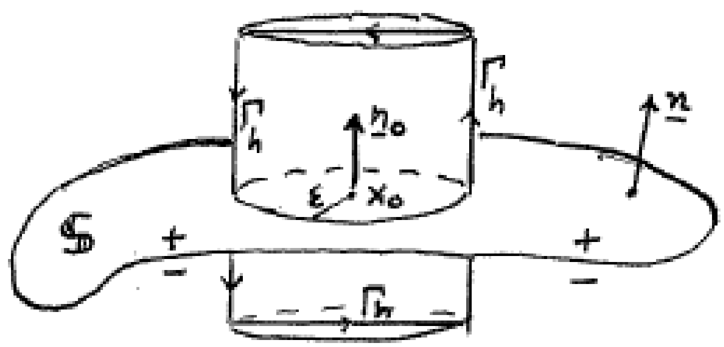

Figure 1.7: The domain Ω_h

We suppose that the total charge contained in Ω_h includes a surface charge distributed over $\mathbb{S}$ with continuous surface density σ. By taking h sufficiently small the total charge contained in Ω_h will be given by

$$Q[\Omega_h] = \int_{\Omega_h} \rho(\boldsymbol{x}, t)\, \boldsymbol{dx} + \int_{\mathbb{I}_o} \sigma(\boldsymbol{y}, t)\, dS_{\boldsymbol{y}}$$

where the function ρ is essentially bounded in Ω_h. Applying (I1) to Ω_h then yields

$$\int_{\partial\Omega_h} \boldsymbol{D}(\boldsymbol{x}, t) \cdot \boldsymbol{n}(\boldsymbol{x})\, dS_{\boldsymbol{x}} = \int_{\Omega_h} \rho(\boldsymbol{x}, t)\, \boldsymbol{dx} + \int_{\mathbb{I}_o} \sigma(\boldsymbol{y}, t) dS_y$$

By taking the limit as $h \to 0$ for fixed ε the integral over Ω_h tends to zero and

$$\lim_{h\to 0} \int_{\partial\Omega_h} \boldsymbol{D}(\boldsymbol{x}, t) \cdot \boldsymbol{n}(\boldsymbol{x})\, dS_{\boldsymbol{x}} = \int_{\mathbb{I}_o} \Big\{ \boldsymbol{D}_-(\boldsymbol{y}, t) \cdot \boldsymbol{n}_-(\boldsymbol{y}) + \boldsymbol{D}_+(\boldsymbol{y}, t) \cdot \boldsymbol{n}_+(\boldsymbol{y}) \Big\}\, dS_{\boldsymbol{y}}$$

where $\boldsymbol{n}_-(\boldsymbol{y}) = -\boldsymbol{n}(\boldsymbol{y})$, $\boldsymbol{n}_+(\boldsymbol{y}) = \boldsymbol{n}(\boldsymbol{y})$. Therefore

$$\int_{\mathbb{I}_o} \Big\{ [\boldsymbol{D}(\boldsymbol{y}, t)]_{\mathbb{S}} \cdot \boldsymbol{n}(\boldsymbol{y}) - \sigma(\boldsymbol{y}, t) \Big\}\, dS_{\boldsymbol{y}} = 0$$

By the mean value theorem there exists a point $\overline{\boldsymbol{y}} \in \mathbb{I}_o$ such that

$$\int_{\mathbb{I}_o} \Big\{ [\boldsymbol{D}(\boldsymbol{x}, t)]_{\mathbb{S}} \cdot \boldsymbol{n}(\boldsymbol{x}) - \sigma(\boldsymbol{x}, t) \Big\}\, dS_{\boldsymbol{x}} = \Big\{ [\boldsymbol{D}(\overline{\boldsymbol{y}}, t)]_{\mathbb{S}} \cdot \boldsymbol{n}(\overline{\boldsymbol{y}}) - \sigma(\overline{\boldsymbol{y}}, t) \Big\} |\mathbb{I}_o| = 0$$

where $|\mathbb{I}_o|$ denotes the area of $\mathbb{I}_o$. Hence $[\boldsymbol{D}(\overline{\boldsymbol{y}},t)]_{\mathbb{S}}\boldsymbol{\cdot}\boldsymbol{n}(\overline{\boldsymbol{y}}) = \sigma(\overline{\boldsymbol{y}},t)$. Since $[\boldsymbol{D}(\boldsymbol{y},t)]_{\mathbb{S}}$ and $\sigma(\boldsymbol{y},t)$ are continuous on $\mathbb{S}$ by hypothesis, letting $\varepsilon \to 0$ yields

$$[\boldsymbol{D}(\boldsymbol{x}_o,t) \ \boldsymbol{\cdot}\, \boldsymbol{n}_o]_{\mathbb{S}} = \sigma(\boldsymbol{x}_o,t)$$

As $[\boldsymbol{D}\boldsymbol{\cdot}\boldsymbol{n}]_{\mathbb{S}} \equiv [\boldsymbol{D}]_{\mathbb{S}}\boldsymbol{\cdot}\boldsymbol{n}$ and $\boldsymbol{x}_o$ is an arbitrary point of $\mathbb{S}$, this proves (R1).

The same argument applied to eq. (I4) for $S = \partial\Omega_h$

$$\int_{\partial\Omega_h} \boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{n}\ dS \ = \ 0$$

proves (R2). Similarly, eq. (I3) for $S = \partial\Omega_h$

$$\int_{\partial\Omega_h} (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t})\boldsymbol{\cdot}\boldsymbol{n}\ dS \ = \ 0$$

yields the equation of conservation of surface charge

$$[\boldsymbol{J}]_{\mathbb{S}}\boldsymbol{\cdot} n + \frac{\partial}{\partial t}[\boldsymbol{D}\boldsymbol{\cdot}\boldsymbol{n}]_{\mathbb{S}} = 0$$

which, owing to (R1), coincides with (R3).

In order to prove (R4), consider the rectangle S_h obtained by intersecting Ω_h with an arbitrary plane containing $\boldsymbol{x}_o$ and $\boldsymbol{n}_o$, and let $\Gamma_h = \partial\Omega_h$ be oriented according to the right-handed screw rule. Applying (I3) to S_h then yields

$$\int_{S_h} (\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t})\boldsymbol{\cdot}\boldsymbol{n}\ dS = \int_{\Gamma_h} \boldsymbol{H}\boldsymbol{\cdot}\boldsymbol{t}\, ds$$

Since $\boldsymbol{J}$ and $\frac{\partial \boldsymbol{D}}{\partial t}$ are bounded by assumption, the surface integral on the left-hand side tends to zero as $h \to 0$. Because the limits $\boldsymbol{H}_{\pm}(\boldsymbol{y},t)$ exist and are continuous by hypothesis, letting also $\varepsilon \to 0$ we find

$$\boldsymbol{H}_{+}(\boldsymbol{x}_o,t)\boldsymbol{\cdot}\boldsymbol{t}_o - \boldsymbol{H}_{-}(\boldsymbol{x}_o,t)\boldsymbol{\cdot}\boldsymbol{t}_o \quad = 0$$

where $\boldsymbol{t}_o$ is the unit vector tangent to $\mathbb{S}$ at the point $\boldsymbol{x}_o$ lying in the plane of S_h. As this plane is arbitrary, and $\boldsymbol{x}_o$ is an arbitrary point on $\mathbb{S}$, the last relation coincides with (R4).

Finally (R5) can be proven by the same argument applied to (I2). This concludes the proof.

Remarks.

1. The surface charge density σ is given by the jump of $\boldsymbol{D}\cdot\boldsymbol{n}$ even if $\epsilon_+ = \epsilon_-$.

2. Eq. (R3) implies that $[\boldsymbol{J}]\cdot\boldsymbol{n} = 0$ at all points of $\mathbb{S}$ where $\partial\sigma/\partial t = 0$. In particular, since $\boldsymbol{J}$ is zero in a dielectric, the boundary condition

$$\boldsymbol{J}\cdot\boldsymbol{n} = 0 \tag{1.41}$$

holds at the interface of a conductor with a dielectric if the surface charge density σ does not depend on t.

3. The matching relation (R5) enables one to define conceptually the electric field inside a dielectric medium by inserting a test charge in an empty narrow cavity parallel to the lines of force of $\boldsymbol{E}$: the field inside the cavity, by force of (R5), coincides with that in the dielectric. Similarly, in order to define $\boldsymbol{D}$ the narrow cavity should be taken orthogonal to the lines of force of $\boldsymbol{E}$ in the dielectric (Exercise 11).

4. If $\boldsymbol{E} = -grad\, u$, the potential u may not be continuous across $\mathbb{S}$, but the jump $[u]_{\mathbb{S}}$ is constant on every connected component $\mathbb{S}_j$ of $\mathbb{S}$. Indeed, (R5) implies

$$[\boldsymbol{n}\wedge grad\, u]_{\mathbb{S}} = \boldsymbol{0}$$

so that the tangential derivatives of u are continuous across $\mathbb{S}$.

Since

$$\boldsymbol{n}\wedge [grad\, u]_{\mathbb{S}} \equiv \boldsymbol{n}\wedge grad\, [u]_{\mathbb{S}}$$

(Exercise 12) this is tantamount to saying that the tangential derivatives of the jump of u vanish on $\mathbb{S}$

$$\boldsymbol{n}\wedge grad\, [u]_{\mathbb{S}} = \boldsymbol{0} \tag{1.42}$$

Therefore if $\mathbb{S}$ is connected $[u]_{\mathbb{S}}$ is constant on $\mathbb{S}$;if $\mathbb{S} = \bigcup_j \mathbb{S}_j$ with $\mathbb{S}_j$ connected, then

$$[u]_{\mathbb{S}} = \mathrm{M}_j \qquad \text{on } \mathbb{S}_j \tag{1.43}$$

where the constants M_j may be different on the different connected components $\mathbb{S}_j$ of $\mathbb{S}$. Conversely, eq. (1.43) obviously implies that $[\boldsymbol{n}\wedge grad\, u]_{\mathbb{S}} = \boldsymbol{0}$.

5. By combining the matching equations (R1)−(R5) with the constitutive relations (C1)−(C3) one can complete the set of matching conditions for all the remaining components of $\boldsymbol{D}$, $\boldsymbol{B}$, $\boldsymbol{H}$, $\boldsymbol{E}$. In particular, (R1) and (C2) imply that the normal component of $\boldsymbol{E}$ satisfies

$$(\epsilon_+ \boldsymbol{E}_+ \cdot \boldsymbol{n} - \epsilon_- \boldsymbol{E}_-) \cdot \boldsymbol{n} \quad \equiv \quad [\epsilon \boldsymbol{E} \cdot \boldsymbol{n}]_{\mathbb{S}} = \ \sigma$$

while (R2) and (C3) show that the normal component of $\boldsymbol{H}$ satisfies

$$(\mu_+ \boldsymbol{H}_+ \cdot \boldsymbol{n} - \mu_- \boldsymbol{H}_-) \cdot \mathbf{n} \quad \equiv \quad [\mu \boldsymbol{H} \cdot \boldsymbol{n}]_{\mathbb{S}} = \ 0$$

These relations can be written in the form

$$\text{(1.44)} \quad [\boldsymbol{E} \cdot \boldsymbol{n}]_{\mathbb{S}} = \frac{\sigma}{\epsilon_+} - \frac{\epsilon_+ - \epsilon_-}{\epsilon_+} \ \boldsymbol{E}_- \cdot \boldsymbol{n} \quad , \quad [\boldsymbol{H} \cdot \boldsymbol{n}]_{\mathbb{S}} = -\frac{\mu_+ - \mu_-}{\mu_+} \ \boldsymbol{H}_- \cdot \boldsymbol{n}$$

where

$$\frac{1}{\epsilon_+} \boldsymbol{E}_- \cdot \boldsymbol{n} \equiv \frac{1}{\epsilon_-} \ \boldsymbol{E}_+ \cdot \boldsymbol{n} \quad , \quad \frac{1}{\mu_+} \boldsymbol{H}_- \cdot \boldsymbol{n} \equiv \frac{1}{\mu_-} \ \boldsymbol{H}_+ \cdot \boldsymbol{n}$$

Since any vector $\boldsymbol{v}$ is determined by its components $\boldsymbol{v} \cdot \boldsymbol{n}$ and $\boldsymbol{v} \wedge \boldsymbol{n}$ via the identity

$$\text{(1.45)} \qquad \boldsymbol{v} \equiv \boldsymbol{v} \cdot \boldsymbol{n} \ \boldsymbol{n} - (\boldsymbol{v} \wedge \boldsymbol{n}) \wedge \boldsymbol{n}$$

it follows that

$$[\boldsymbol{v}]_{\mathbb{S}} = [\boldsymbol{v} \cdot \boldsymbol{n}]_{\mathbb{S}} \boldsymbol{n} - [\boldsymbol{v} \wedge \boldsymbol{n}]_{\mathbb{S}} \wedge \boldsymbol{n}$$

The complete matching relations for the field vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ read then

$$\text{(1.46)} \qquad [\boldsymbol{E}]_{\mathbb{S}} = [\boldsymbol{E} \cdot \boldsymbol{n}]_{\mathbb{S}} \boldsymbol{n} \quad , \quad [\boldsymbol{H}]_{\mathbb{S}} = [\boldsymbol{H} \cdot \boldsymbol{n}]_{\mathbb{S}} \boldsymbol{n}$$

where $[\boldsymbol{E} \cdot \boldsymbol{n}]_{\mathbb{S}}$ and $[\boldsymbol{H} \cdot \boldsymbol{n}]_{\mathbb{S}}$ are given in (1.44). Similarly the complete matching relations for the vectors $\boldsymbol{D}$ and $\boldsymbol{B}$ are
(1.47)

$$[\boldsymbol{D}]_{\mathbb{S}} = \sigma \boldsymbol{n} - \frac{\epsilon_+ - \epsilon_-}{\epsilon_-} (\boldsymbol{D}_- \wedge \boldsymbol{n}) \wedge \boldsymbol{n} \quad , \quad [\boldsymbol{B}]_{\mathbb{S}} = -\frac{\mu_+ - \mu_-}{\mu_-} (\boldsymbol{B}_- \wedge \boldsymbol{n}) \wedge \boldsymbol{n}$$

where

$$\frac{1}{\epsilon_-} \boldsymbol{D}_- \wedge \boldsymbol{n} \equiv \frac{1}{\epsilon_+} \boldsymbol{D}_+ \wedge \boldsymbol{n} \quad , \quad \frac{1}{\mu_-} \boldsymbol{B}_- \wedge \boldsymbol{n} \equiv \frac{1}{\mu_+} \ \boldsymbol{B}_+ \wedge \boldsymbol{n}$$

These relations show that

(i) $\boldsymbol{E}$ and $\boldsymbol{D}$ are continuous if and only if $\epsilon_+ = \epsilon_-$ and $\sigma = 0$;

(ii) $\boldsymbol{B}$ and $\boldsymbol{H}$ are continuous if and only if $\mu_+ = \mu_-$

Two consequences of the matching relations are contained in the next propositions.

Proposition 1.4.1 *Coulomb's law (1.38) extends to a homogeneous dielectric ball of arbitrary radius centered at the point charge Q and surrounded by another homogeneous unbounded dielectric medium.*

Proof. Let the charge Q be concentrated at the point $\boldsymbol{x}_o$ and let Ω_- and Ω_+ coincide with the interior and exterior of the ball $B_R := \{\mathbf{x} \in \mathbb{R}^3 \colon |\boldsymbol{x} - \boldsymbol{x}_o| \leq R\}$, respectively, so that $\mathbb{S}$ is the sphere $|\boldsymbol{x} - \boldsymbol{x}_o| = R$. If we take $\boldsymbol{D}(\boldsymbol{x}) = D(r)\boldsymbol{n}$ as in Proposition 1.3.4 and Proposition 1.3.5, where $\boldsymbol{n} = (\boldsymbol{x} - \boldsymbol{x}_o)/|\boldsymbol{x} - \boldsymbol{x}_o|$ is the outer normal to $\mathbb{S}$, the matching relation (1.47) shows that $\boldsymbol{D}$ is continuous across $\mathbb{S}$:

$$\boldsymbol{D}_+ - \boldsymbol{D}_- = \mathbf{0}$$

This relation and the Gauss law are satisfied by the Coulomb law (1.5)

$$\boldsymbol{D}(\boldsymbol{x}) = \frac{Q}{4\pi} \frac{\boldsymbol{x} - \boldsymbol{x}_o}{r^3} \qquad (r = |x - x_o| > 0)$$

The electric field

$$\boldsymbol{E}(\boldsymbol{x}) = \begin{cases} \frac{Q}{4\pi\epsilon_-} \frac{\boldsymbol{x} - \boldsymbol{x}_o}{r^3} & \text{for } 0 < r < R \\ \frac{Q}{4\pi\epsilon_+} \frac{\boldsymbol{x} - \boldsymbol{x}_o}{r^3} & \text{for } r > R \end{cases}$$

has a jump discontinuity for $|\boldsymbol{x} - \boldsymbol{x}_o| = R$, in accordance with eq. (1.46). The electric potential, vanishing at infinity and continuous across $\mathbb{S}$, is given by

$$u(\boldsymbol{x}) = \begin{cases} \frac{Q}{4\pi\epsilon_- r} & \text{for } 0 < r < R \\ \frac{Q}{4\pi\epsilon_+ r} + \frac{Q}{4\pi R}\left(\frac{1}{\epsilon_-} - \frac{1}{\epsilon_+}\right) & \text{for } r > R \end{cases}$$

Proposition 1.4.2 *On the surface $\mathbb{S}$ of a perfect conductor the tangential component of $\boldsymbol{E}$ is zero:*

$$\boldsymbol{E} \wedge \boldsymbol{n} = \mathbf{0} \Leftrightarrow \boldsymbol{E}_{tang} = \mathbf{0}$$

Proof. $\gamma_- = +\infty$ implies $\boldsymbol{E}_- = \boldsymbol{0}$ in Ω_- by force of Proposition 1.3.2 , hence $\boldsymbol{E}_- \wedge \boldsymbol{n} = \boldsymbol{0}$ on $\mathbb{S}$. Since $\boldsymbol{E} \wedge \boldsymbol{n}$ is continuous because of (R5), it follows that $\boldsymbol{E}_{tang} = -(\boldsymbol{E} \wedge \boldsymbol{n}) \wedge \boldsymbol{n} = \boldsymbol{0}$

A further important consequence of the matching relations and of the linear constitutive laws (C2), (C3) is the refraction of the lines of force of the electric or magnetic field at the boundary between two different media.

Proposition 1.4.3 *At an uncharged boundary surface $\mathbb{S}$ between two non-ferromagnetic media with permittivities ϵ_-, ϵ_+ and permeabilities μ_-, μ_+, respectively, the lines of force of* **E, D** *and* **H, B** *refract according to the laws*

$$\epsilon_- \, tan\alpha_+ = \epsilon_+ \, tan\alpha_- \quad , \quad \mu_- \, tan\alpha_+ = \mu_+ \, tan\alpha_-$$

respectively, where α_-, α_+ are the angles between the vector $\boldsymbol{E}$ or $\boldsymbol{H}$ and the normal $\boldsymbol{n}$ in the two dielectric media. Similarly, if the two media are conductors with conductivities γ_-, γ_+ and the surface charge density σ on $\mathbb{S}$ is time-independent, the law of refraction for the lines of current of $\boldsymbol{J}$ (hence also of $\boldsymbol{E}$) can be written as

$$\gamma_- \, tan\alpha_+ = \gamma_+ \, tan\alpha_-$$

Proof. Equations (1.45) imply that $(\boldsymbol{E}_-, \boldsymbol{E}_+, \boldsymbol{n})$ and $(\boldsymbol{H}_-, \boldsymbol{H}_+, \boldsymbol{n})$ are coplanar. Let $\sigma = 0$ on $\mathbb{S}$. Eqs. (R1), (R5) and (C2) yield then

$$\epsilon_- |\boldsymbol{E}_-| cos\alpha_- = \epsilon_+ |\boldsymbol{E}_+| cos\alpha_+ \quad , \quad |\boldsymbol{E}_-| sin\alpha_- = |\boldsymbol{E}_+| sin\alpha_+$$

and the refraction law for $\boldsymbol{E}$ follows immediately from these relations. As $\boldsymbol{D}$ is parallel to $\boldsymbol{E}$ because of (C2), the same law holds also for $\boldsymbol{D}$. Similarly, eqs. (R2), (R4) and (C3) yield

$$\mu_- |\boldsymbol{H}_-| cos\alpha_- = \mu_+ |\boldsymbol{H}_+| cos\alpha_+ \quad , \quad |\boldsymbol{H}_-| sin\alpha_- = |\boldsymbol{H}_+| sin\alpha_+$$

and we obtain the refraction law for $\boldsymbol{H}$. Since $\boldsymbol{B}$ is parallel to $\boldsymbol{H}$ from eq. (C3), the same law holds also for $\boldsymbol{B}$. If $\mathbb{S}$ separates two conductors and $\partial\sigma/\partial t = 0$ on $\mathbb{S}$, the matching relation (R3) becomes

$$[\boldsymbol{J}]_{\mathbb{S}} \cdot \boldsymbol{n} = 0$$

that is, $\boldsymbol{J} \cdot \boldsymbol{n}$ is continuous. Since $\boldsymbol{J} = \gamma \boldsymbol{E}$, we have the matching relations

$$\gamma_- \boldsymbol{E}_- \cdot \boldsymbol{n} = \gamma_+ \boldsymbol{E}_+ \cdot \boldsymbol{n} \ , \quad [\boldsymbol{E}]_{\mathbb{S}} \wedge \boldsymbol{n} = \boldsymbol{0}$$

and by the same argument as before we find the refraction law $\gamma_- \, tan\alpha_+ = \gamma_+ \, tan\alpha_-$.

Note that in the case when $\mathbb{S}$ separates two conductors, the last equation implies that

$$\frac{\gamma_-}{\epsilon_-}\boldsymbol{D}_- \cdot \boldsymbol{n} = \frac{\gamma_+}{\epsilon_+}\boldsymbol{D}_+ \cdot \boldsymbol{n}$$

so that $\boldsymbol{D} \cdot \boldsymbol{n}$ is discontinuous across $\mathbb{S}$, and hence $\sigma \neq 0$, unless

$$\frac{\gamma_-}{\epsilon_-} = \frac{\gamma_+}{\epsilon_+}$$

In this particular case $\sigma \equiv 0$ and

$$\frac{tan\alpha_+}{tan\alpha_-} = \frac{\gamma_+}{\gamma_-} \equiv \frac{\epsilon_+}{\epsilon_-}$$

according to the fact that the two refraction laws for $\boldsymbol{E}$ and $\boldsymbol{J}$ must coincide.

Corollary 1.4.4 *The separation surface* $\mathbb{S}$ *between two conductors with* $\gamma_+ \ll \gamma_-$ *is equipotential in steady conditions.*

Proof. The assumption implies that $\boldsymbol{E}_- \cong \boldsymbol{0}$, $\alpha_+ \simeq 0$, and $\boldsymbol{E}_+$ is orthogonal to $\mathbb{S}$. Since $curl\boldsymbol{E} = \boldsymbol{0}$ in steady conditions, $\boldsymbol{E} = -grad\, u$ is orthogonal to $\mathbb{S}$, which means that u is constant along $\mathbb{S}$.

1.5 Energy balance and uniqueness theorem

1.5.1 Electromagnetic energy and Poynting vector.

We proceed to define the energy of the electromagnetic field in a region of space and to derive an energy balance equation. Suppose that a normal domain Ω of $\mathbb{R}^3$ is filled by a material medium having smooth coefficients $\epsilon(\boldsymbol{x}), \mu(\boldsymbol{x}), \gamma(\boldsymbol{x})$. More precisely, suppose that:

(**H1**) $\epsilon(\boldsymbol{x}), \mu(\boldsymbol{x}), \gamma(\boldsymbol{x})$ are continuous and bounded functions in Ω, with

$$\epsilon(x) \geq \epsilon_o > 0 \quad , \quad \mu(x) \geq \overline{\mu} > 0, \quad \gamma(\boldsymbol{x}) \geq 0$$

for all $\boldsymbol{x} \in \Omega$, and

(H2) $\boldsymbol{E}(\boldsymbol{x},t), \boldsymbol{H}(\boldsymbol{x},t)$ are of class $C^{\text{o}}(\overline{\Omega} \times [0,+\infty)) \cap C^1(\overline{\Omega} \times (0,+\infty))$.

The energy theorem can be obtained by manipulating the Maxwell equations (M1) and (M2) as follows. By multiplying scalarly eq. (M1) by **H**, equation (M2) by $\boldsymbol{E}$, and then adding up we obtain

$$\boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{E} \cdot curl \boldsymbol{H} - \boldsymbol{H} \cdot curl \boldsymbol{E} - \boldsymbol{E} \cdot \boldsymbol{J}$$

The constitutive relations (C1)−(C3) imply $\boldsymbol{E} \cdot \boldsymbol{J} = \gamma |\boldsymbol{E}|^2$, and

$$\boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} = \epsilon \boldsymbol{E} \cdot \frac{\partial \boldsymbol{E}}{\partial t} = \frac{\partial}{\partial t}(\frac{1}{2}\epsilon |\boldsymbol{E}|^2)$$

$$\boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} = \mu \boldsymbol{H} \cdot \frac{\partial \boldsymbol{H}}{\partial t} = \frac{\partial}{\partial t}(\frac{1}{2}\mu |\boldsymbol{H}|^2)$$

Moreover, a well-known vector identity yields

$$\boldsymbol{E} \cdot curl \boldsymbol{H} - \boldsymbol{H} \cdot curl \boldsymbol{E} \equiv - div(\boldsymbol{E} \wedge \boldsymbol{H}) \tag{1.48}$$

Taking all these relations into account we find the equation

$$\frac{\partial W}{\partial t} + div\, \mathbf{S} = -P_J \tag{1.49}$$

where

$$W = W_e + W_m \equiv \frac{1}{2}(\epsilon |\boldsymbol{E}|^2 + \mu |\boldsymbol{H}|^2)$$

is interpreted as the total energy density of the electromagnetic field,

$$W_e := \frac{1}{2} \boldsymbol{E} \cdot \boldsymbol{D} \equiv \frac{1}{2}\epsilon |\boldsymbol{E}|^2$$

is the electrical energy density,

$$W_m = \frac{1}{2} \boldsymbol{H} \cdot \boldsymbol{B} \equiv \frac{1}{2}\mu |\boldsymbol{H}|^2$$

is the magnetic energy density, and

$$\boldsymbol{S} := \boldsymbol{E} \wedge \boldsymbol{H}$$

is called the Poynting vector. Moreover,

$$P_J = \boldsymbol{E} \cdot \boldsymbol{J} \; = \gamma|\boldsymbol{E}|^2$$

is the work per unit time and unit volume done by the electric field over the current $\boldsymbol{J}$. This work is zero for dielectrics, where $\gamma = 0,$ whereas in a conductor, where $0 < \gamma < +\infty$, it can be equivalently expressed in terms of the resistivity γ^{-1} and of the current $\boldsymbol{J}$ as

$$P_J \equiv \gamma^{-1}|\boldsymbol{J}|^2$$

It is well-known that, for a wire of finite cross-section A and length l the power dissipated into heat due to the "Joule effect", i.e. the rate at which heat is evolved in a resistance, is given by

$$P = \mathrm{V}I \equiv \mathcal{R}I^2$$

where $V = I\mathcal{R}$ is the e.m.f., $I \; = A|\boldsymbol{J}|$ is the current, and $\mathcal{R}$ the resistance of the wire, given by eq. (1.35). It is then easy to see that $P_J = \mathcal{R}I^2/lA$, where lA is the volume of the wire (Exercise 13).

Thus P_J can be interpreted as the power per unit volume dissipated into heat by the Joule effect. It remains to clarify the meaning of the Poynting vector $\boldsymbol{E}\wedge\boldsymbol{H}$.

The energy of the electromagnetic field is thought of as being distributed through the region of space where the field is present by means of the energy density W. Thus the total (local) electromagnetic energy in the domain Ω is defined by

$$\mathcal{E}(t) = \mathcal{E}_\Omega(t) := \int_\Omega (\frac{1}{2}\epsilon|\boldsymbol{E}(\boldsymbol{x},t)|^2 + \; \frac{1}{2}\mu|\boldsymbol{H}(\boldsymbol{x},t)|^2)\, \boldsymbol{dx} \tag{1.50}$$

and the energy balance equation is obtained by integrating eq. (1.49) over a normal domain Ω and applying the divergence theorem: if $\boldsymbol{n}$ is the outer normal to $\partial\Omega$ we obtain

$$\frac{d\mathcal{E}}{dt} = -\int_\Omega \gamma|\boldsymbol{E}|^2\, \boldsymbol{dx} - \int_{\partial\Omega} \boldsymbol{S} \cdot \boldsymbol{n}\, dS \tag{1.51}$$

As $\epsilon > 0,\ \mu > 0,\ \mathcal{E}(t)$ is non-negative for every t. The integral involving S on the r.h.s. is interpreted as the outgoing power flux through the boundary surface $\partial\Omega$, and so the Poynting vector $\boldsymbol{S}=\boldsymbol{E}\wedge\boldsymbol{H}$ represents the power flux vector (power per unit surface). According to eq. (1.51) the local energy $\mathcal{E}_\Omega(t)$ changes solely because of the power dissipated into heat in the domain Ω by the Joule effect, given by the volume integral on the r.h.s. of eq. (1.51), and of the outgoing power flux through the boundary $\partial\Omega$, given by the surface integral of $\boldsymbol{S}\cdot\boldsymbol{n}$. The Poynting vector is orthogonal to both $\boldsymbol{E}$ and $\boldsymbol{H}$, and in any Ohmic conductor, where $\boldsymbol{E}=\gamma^{-1}\boldsymbol{J}, 0<\gamma<+\infty,$ the power flux vector $\boldsymbol{S}$ is always orthogonal to the current $\boldsymbol{J}$. On the other hand, dielectrics cannot conduct electricity but, as we shall see, they do carry power under the form of radiating electromagnetic waves, and the Poynting vector is then called the radiation vector.

In certain cases $\boldsymbol{S}\cdot\boldsymbol{n}=0$ and the power flux vanishes.

Proposition 1.5.1 *Let $\mathbb{S}$ denote a separation surface between two different media, as in §1.4. Then the power flux density $\boldsymbol{S}\cdot\boldsymbol{n}$ is continuous across $\mathbb{S}$:*

$$[\boldsymbol{S}\cdot\boldsymbol{n}]_{\mathbb{S}} := \boldsymbol{S}_+\cdot\boldsymbol{n} - \boldsymbol{S}_-\cdot\boldsymbol{n} = 0 \tag{1.52}$$

(ii) If $\boldsymbol{E}\wedge\boldsymbol{n}=\boldsymbol{0}$ or $\boldsymbol{H}\wedge\boldsymbol{n}=\boldsymbol{0}$ at some point $\boldsymbol{y}\in\partial\Omega$, then $\boldsymbol{S}\cdot\boldsymbol{n}=0$ there.

In particular the power flux across $\partial\Omega$ vanishes if the tangential component of either $\boldsymbol{E}$ or $\boldsymbol{H}$ vanishes on $\partial\Omega$.

Proof. We have $\boldsymbol{S}\cdot\boldsymbol{n}=\boldsymbol{E}\wedge\boldsymbol{H}\cdot\boldsymbol{n}\equiv\boldsymbol{H}\wedge\boldsymbol{n}\cdot\boldsymbol{E}\equiv-\boldsymbol{E}\wedge\boldsymbol{n}\cdot\boldsymbol{H}$, hence

$$\boldsymbol{S}\cdot\boldsymbol{n} = \boldsymbol{E}_{tang}\wedge\boldsymbol{H}_{tang}\cdot\boldsymbol{n} \tag{1.53}$$

and the assertions follow immediately from eqs. (1.40), (R4), (R5).

Equation (1.52) shows that at an interface between two adjoining material bodies the power flux going out of one body equals the power flux going into the other.

Proposition 1.5.2 *Through the boundary of a perfect conductor ($\gamma=+\infty$) no power flux is possible.*

Proof. From Proposition 1.5.1 and Proposition 1.4.2 it follows that $\boldsymbol{S}\cdot\boldsymbol{n}=0$.

Proposition 1.5.3 *The energy balance equation (1.51) holds also if*

(i) $\gamma(\boldsymbol{x})$, $\epsilon(\boldsymbol{x})$, $\mu(\boldsymbol{x})$ *are bounded and piecewise smooth, with jump discontinuities across a surface* $\mathbb{S}$ *in* Ω

(ii) Ω *is an unbounded domain external to a closed connected surface* S *and* $\boldsymbol{E}$, $\boldsymbol{H}$ *satisfy the asymptotic condition at infinity*

$$\boldsymbol{E}(\boldsymbol{x},t) = O(|\boldsymbol{x}|^{-2}), \ \boldsymbol{H}(\boldsymbol{x},t) = O(|\boldsymbol{x}|^{-2}) \tag{1.54}$$

as $|\boldsymbol{x}| \to \infty$ *(uniformly with respect to t and to direction)* [9].

(iii) $\Omega = \mathbb{R}^3$, $\gamma(\boldsymbol{x})$, $\epsilon(\boldsymbol{x})$, $\mu(\boldsymbol{x})$ *are bounded and piecewise smooth, and* $\boldsymbol{E}$, $\boldsymbol{H}$ *satisfy (1.54).*

Proof. (i) Let $\mathbb{S}$ be the separation surface between two different media Ω_-, Ω_+ and let

$$\Omega = \Omega_- \cup \Omega_+, \quad \Omega_- \cap \Omega_+ = \emptyset, \quad \partial\Omega_- = \Sigma_- \cup \mathbb{S}, \quad \partial\Omega_+ = \Sigma_+ \cup \mathbb{S}$$

with the normal $\boldsymbol{n}$ to $\mathbb{S}$ oriented from Ω_- to Ω_+. Since eq. (1.51) holds separately for Ω_- and Ω_+ , we have the two equations

$$\frac{d\mathcal{E}_{\Omega_-}(t)}{dt} = -\int_{\Omega_-} \gamma|\boldsymbol{E}|^2 \, \boldsymbol{dx} - \int_{\Sigma_-} \boldsymbol{S} \cdot \boldsymbol{n} \, dS - \int_{\mathbb{S}} \boldsymbol{S}_- \cdot \boldsymbol{n} \, dS$$
$$\frac{d\mathcal{E}_{\Omega_+}(t)}{dt} = -\int_{\Omega_+} \gamma|\boldsymbol{E}|^2 \, \boldsymbol{dx} - \int_{\Sigma_+} \boldsymbol{S} \cdot \boldsymbol{n} \, dS + \int_{\mathbb{S}} \boldsymbol{S}_+ \cdot \mathbf{n} \, dS$$

The energy $\mathcal{E}_\Omega$ is an additive set function, $\mathcal{E}_{\Omega_-} + \mathcal{E}_{\Omega_+} = \mathcal{E}_\Omega$, and $\boldsymbol{S}_- \cdot \boldsymbol{n} = \boldsymbol{S}_+ \cdot \boldsymbol{n}$ by Proposition 1.5.1. Adding up the two preceding equations and taking (1.54) into account yields then the energy balance equation (1.51) for $\Omega_- \cup \Omega_+ = \Omega$. (ii) Let $\Omega = \mathbb{R}^3 \backslash \overline{\mathcal{Q}}$, where $\mathcal{Q}$ is a normal domain in $\mathbb{R}^3$ with connected boundary $S = \partial\mathcal{Q}$. Consider the truncated normal domain $\Omega_R := \{\boldsymbol{x} \in \Omega : |\boldsymbol{x}| < R\}$. The energy balance equation for Ω_R reads then

$$\frac{d\mathcal{E}_{\Omega_R}}{dt} = -\int_{\Omega_R} \gamma|\boldsymbol{E}|^2 \, \boldsymbol{dx} - \int_{\partial\Omega_R} \boldsymbol{E} \wedge \boldsymbol{H} \cdot \boldsymbol{n} \, dS$$

[9] i.e. there exist constants $M, R > 0$ independent of t such that $||\mathbf{x}|^2\mathbf{E}(\mathbf{x},t)| < M$ for all $|\mathbf{x}| > R$

where $\partial\Omega_R = \partial\mathcal{Q} \cup \{|\boldsymbol{x}| = R\}$ and $\boldsymbol{n}$ denotes the outer normal to $\partial\Omega_R$. Our assumptions imply that $\boldsymbol{E}$,$\boldsymbol{H}$ are in $L^2(\Omega)$ for every t and that $\boldsymbol{E}\wedge\boldsymbol{H}= O(|\boldsymbol{x}|^{-4})$ as $|\boldsymbol{x}| \to \infty$. Therefore, letting $R \to \infty$ the Poynting vector integral tends to zero and we obtain the energy balance equation for Ω

$$\frac{d\mathcal{E}_\Omega}{dt} = -\int_\Omega \gamma|\boldsymbol{E}|^2\, \boldsymbol{dx} - \int_S \boldsymbol{E} \wedge \boldsymbol{H} \cdot \boldsymbol{n}\, dS$$

where the normal $\boldsymbol{n}$ on S is exterior to Ω, i.e. is interior to $\mathcal{Q} = \mathbb{R}^3\backslash\overline{\Omega}$.

(*iii*) If $\Omega = \mathbb{R}^3$ the Poynting power flux integral over $S = \partial\mathcal{Q}$ disappears and we obtain

$$\frac{d\mathcal{E}_{\mathbb{R}^3}}{dt} = -\int_{\mathbb{R}^3} \gamma|\boldsymbol{E}|^2\, \boldsymbol{dx}$$

Condition (1.54) implies that the power flux at infinity is zero.

Proposition 1.5.4 *The electromagnetic energy $\mathcal{E}_\Omega(t)$ is a non-increasing function of t if $\gamma(\boldsymbol{x})$, $\epsilon(x)$, $\mu(\boldsymbol{x})$ are bounded and piecewise smooth and Ω satisfies either $(i), (ii)$ or (iii) :*

(i) Ω is a normal domain with $\boldsymbol{E}\wedge\boldsymbol{n}= \boldsymbol{0}$ or $\boldsymbol{H}\wedge\boldsymbol{n}= \boldsymbol{0}$ on $\partial\Omega$

(ii) $\Omega = \mathbb{R}^3\backslash\overline{\mathcal{Q}}$, where $\mathcal{Q}$ is a normal domain in $\mathbb{R}^3$ with connected boundary $\partial\mathcal{Q}$ on which $\boldsymbol{E}\wedge\boldsymbol{n}= \boldsymbol{0}$ or $\boldsymbol{H}\wedge\boldsymbol{n}= \boldsymbol{0}$, and ***E, H*** *satisfy condition (1.54) at infinity*

(iii) $\Omega = \mathbb{R}^3$ and ***E, H*** *satisfy (1.54) in $\mathbb{R}^3$.*

More precisely, in all the three cases $\mathcal{E}_\Omega(t)$ is constant in a dielectric material, where $\gamma = 0$, and is dissipated into Joule heat in a conductor, where $0 < \gamma < +\infty$.

Proof. By force of Proposition 1.5.1 and Proposition 1.5.3, in all the three cases $\boldsymbol{S}\cdot\boldsymbol{n}= 0$ and the energy estimate reads

$$\frac{d\mathcal{E}_\Omega}{dt} = -\int_\Omega \gamma(\boldsymbol{x})|\boldsymbol{E}|^2\, \boldsymbol{dx} \tag{1.55}$$

so that $d\mathcal{E}_\Omega/dt \equiv 0$ if $\gamma(\boldsymbol{x}) \equiv 0$, $d\mathcal{E}_\Omega/dt < 0$ if $\gamma(\boldsymbol{x}) > 0$ for $\boldsymbol{x}\in \Omega$.

It is worthwhile to remark that the electric field generated by an isolated point source or dipole has infinite local energy (Exercise 14).

1.5.2 Uniqueness theorem.

The energy balance equation is also relevant mathematically since it enables us to prove uniqueness of the solution ($\boldsymbol{E}$,$\boldsymbol{H}$) to the Initial-Boundary Value (I-BV) problem for the Maxwell equations. The argument of proof, called energy method, is based essentially on the fact that for $\epsilon > 0$, $\mu > 0$ (i.e. excluding superconducting bodies) the energy function $\mathcal{E}_\Omega(t)$ represents the square of a "weighted $L^2(\Omega)-$norm" for the pair of vector valued functions ($\boldsymbol{E}$,$\boldsymbol{H}$) at time t.

Theorem 1.5.5 *(uniqueness for the I-BV problem) Let $\epsilon(\boldsymbol{x}), \mu(\boldsymbol{x})$, $\gamma(\boldsymbol{x})$ be continuous and bounded functions in a normal domain Ω of $\mathbb{R}^3$, with*

$$\epsilon(x) \geq \epsilon_o > 0 \quad , \quad \mu(\boldsymbol{x}) \geq \overline{\mu} > 0, \quad \gamma(x) \geq 0$$

for all $\boldsymbol{x} \in \Omega$. There exists at most one solution $\boldsymbol{E}(\boldsymbol{x},t), \boldsymbol{H}(\boldsymbol{x},t)$ of class $C^o(\overline{\Omega} \times [0,+\infty)) \cap C^1(\overline{\Omega} \times (0,+\infty))$ of the Maxwell equations

$$\mu \frac{\partial \boldsymbol{H}}{\partial t} = -curl \boldsymbol{E} \, , \quad \epsilon \frac{\partial \boldsymbol{E}}{\partial t} = curl \boldsymbol{H} - \gamma \boldsymbol{E} \qquad (\boldsymbol{x} \in \Omega, \, t > 0) \tag{1.56}$$

with initial data

$$\boldsymbol{E}(\boldsymbol{x},0) = \boldsymbol{E}_o(\boldsymbol{x}) \quad , \quad \boldsymbol{H}(\boldsymbol{x},0) = \boldsymbol{H}_o(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \Omega)$$

and with the boundary condition for $\boldsymbol{E} \wedge \boldsymbol{n}$

$$\boldsymbol{E}(\boldsymbol{x},t) \wedge \boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{u}_o(\boldsymbol{x},t) \qquad (\boldsymbol{x} \in \partial\Omega, \, t > 0)$$

or for $\boldsymbol{H} \wedge \boldsymbol{n}$

$$\boldsymbol{H}(\boldsymbol{x},t) \wedge \boldsymbol{n}(\boldsymbol{x}) = v_o(\boldsymbol{x},t) \qquad (\boldsymbol{x} \in \partial\Omega, \, t > 0)$$

where $\boldsymbol{E}_o(\boldsymbol{x}), \boldsymbol{H}_o(\boldsymbol{x}), \boldsymbol{u}_o(\boldsymbol{x},t)$ and $\mathbf{v}_o(\boldsymbol{x},t)$ are arbitrary vector functions.

Proof. Suppose ($\boldsymbol{E}_1$,$\boldsymbol{H}_1$), ($\boldsymbol{E}_2$,$\boldsymbol{H}_2$) are two solutions corresponding to the same set of initial and boundary data. Let

$$\boldsymbol{e} := \boldsymbol{E}_1 - \boldsymbol{E}_2 \qquad \mathbf{h} := \boldsymbol{H}_1 - \boldsymbol{H}_2$$

denote the difference field, and

$$w(\boldsymbol{x},t) := \frac{1}{2}(\epsilon|\boldsymbol{e}|^2 + \mu|\boldsymbol{h}|^2) \quad , \quad \mathcal{E}_o(t) = \int_\Omega w(\boldsymbol{x},t)\,\boldsymbol{dx}$$

the corresponding energy. We have to prove that the two solutions coincide, i.e. that $\boldsymbol{e}(\boldsymbol{x},t) \equiv \boldsymbol{h}(\boldsymbol{x},t) \equiv 0$. By linearity, $\boldsymbol{e}$ and $\boldsymbol{h}$ satisfy the Maxwell system (1.56) with homogeneous initial an boundary data

$$\boldsymbol{e}(\boldsymbol{x},0) = \boldsymbol{h}(\boldsymbol{x},0) = 0 \quad \text{in } \Omega \qquad , \qquad \boldsymbol{e}\wedge\boldsymbol{n} = 0 \;\; or \boldsymbol{h}\wedge\boldsymbol{n} = 0 \quad \text{on } \partial\Omega$$

so that $w(\boldsymbol{x},0) \equiv 0$ in $\Omega\,, \mathcal{E}_o(0) = 0$. By Proposition 1.5.1 the Poynting vector $\boldsymbol{s}{=}\boldsymbol{e}\wedge\boldsymbol{n}$ satisfies $\boldsymbol{s}\cdot\boldsymbol{n}{=}\,0$ on $\partial\Omega$ (for every $t>0$), and so the energy balance equation for $\mathcal{E}_o(t)$ reads

$$\frac{d\mathcal{E}_o}{dt} = -\int_\Omega \gamma|\boldsymbol{e}|^2\,\boldsymbol{dx} \;\; \leq 0$$

The finite increment theorem applied to the function $\mathcal{E}_o(t)$, continuous and non negative for $t \geq 0$, and zero initially, yields

$$0 \leq \; \mathcal{E}_o(t) = \; \mathcal{E}_o(0) + t\frac{d\mathcal{E}_o}{dt}\Big|_{t=\tau} \; \leq \; \mathcal{E}_o(0) = 0 \qquad , \qquad 0 < \tau < t$$

It follows that $\mathcal{E}_o(t) \equiv 0$ for every $t \geq 0$. But $\mathcal{E}_o(t)$ is a weighted L^2-norm for $(\boldsymbol{e},\boldsymbol{h})$, hence $\boldsymbol{e}(\boldsymbol{x},t) \equiv \boldsymbol{h}(\boldsymbol{x},t) \equiv 0$ for every $\boldsymbol{x} \in \Omega, t \geq 0$.

If the boundary $\partial\Omega$ is a perfectly conducting surface or if it is grounded, i.e. maintained at the constant potential $u = 0$, then Proposition 1.4.2 or the relation $\boldsymbol{n}\wedge grad\, u = 0$ imply that $\boldsymbol{E}\wedge\boldsymbol{n} \equiv \boldsymbol{0}$ on $\partial\Omega$ and the uniqueness theorem shows that Ω is shielded off in the sense of the following

Corollary 1.5.6 *(Faraday cage). Suppose* $\boldsymbol{E}\wedge\boldsymbol{n} \equiv \boldsymbol{0}$ *on* $\partial\Omega$ *and* $\boldsymbol{E}(\boldsymbol{x},0) \equiv \boldsymbol{H}(\boldsymbol{x},0) \equiv \boldsymbol{0}$. *Then* $\boldsymbol{E}(\boldsymbol{x},t) \equiv \boldsymbol{H}(\boldsymbol{x},t) \equiv \boldsymbol{0}$ *in* Ω *for every* $t \geq 0$.

Proof. Under the stated assumptions all initial and boundary data vanish and by the previous uniqueness theorem the electromagnetic field inside the domain Ω stays equal to zero, irrespective of what happens outside.

Remarks.

1. The initial data can be assigned for an arbitrary initial time $t = t_o$, since the Maxwell equations are invariant with respect to time translations $t' = t + t_o$.

2. By force of Proposition 1.5.3, the theorem still holds if $\epsilon(\boldsymbol{x}), \mu(\boldsymbol{x})$, $\gamma(\boldsymbol{x})$ are bounded and piecewise continuous with jump discontinuities across one or more smooth surfaces, as it happens in the presence of two or more different materials in Ω.

3. Similarly, Proposition 1.5.4 implies that the uniqueness theorem can be extended to unbounded exterior domains $\Omega = \mathbb{R}^3 \backslash \overline{\mathcal{Q}}$ by adding the asymptotic condition at infinity (1.54) as in Proposition 1.5.3.

1.6 Stationary Maxwell equations

If all the field vectors ***E, D, B, H, J*** are independent of time, the Maxwell equations (M1)-(M5) of §3.1 take the stationary (or steady) form

$$curl\boldsymbol{E}(\boldsymbol{x}) = 0 \quad , \qquad div\boldsymbol{D}(\boldsymbol{x}) = \rho(\boldsymbol{x})$$

and

$$curl\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x}) \; , \; div\boldsymbol{J}(\boldsymbol{x}) = 0 \; , \quad div\boldsymbol{B}(\boldsymbol{x}) = 0$$

The solutions will be called stationary (or steady) fields.

For simplicity we restrict our attention to homogeneous non-magnetic media, characterized by the linear constituitive relations $\boldsymbol{D}= \epsilon\boldsymbol{E}$, $\boldsymbol{B}= \mu\boldsymbol{H}$ and Ohm's law $\boldsymbol{J}= \gamma\boldsymbol{E}$ of §1.3, with $\epsilon \geq \epsilon_o$, $\mu > 0$ and $\gamma \geq 0$ piecewise constant in $\mathbb{R}^3$, that is, constant inside every material medium. We recall that for non-magnetic media we can assume $\mu = \mu_o$ everywhere. The stationary electric and magnetic fields ***E, H*** then satisfy the two systems of equations

$$curl\boldsymbol{E}(\boldsymbol{x}) = 0 \quad , \qquad div\boldsymbol{E}(\boldsymbol{x}) = \rho(\boldsymbol{x})/\epsilon \tag{1.57}$$

and

$$curl\boldsymbol{H}(\boldsymbol{x}) = \gamma\boldsymbol{E}(\boldsymbol{x}) \quad , \qquad div\boldsymbol{H}(\boldsymbol{x}) = 0 \tag{1.58}$$

at the interior of each homogeneous medium, with $\gamma\boldsymbol{E}(\boldsymbol{x}) \equiv \boldsymbol{J}$ satisfying the equations

$$div\boldsymbol{J} \equiv \gamma\, div\boldsymbol{E}(\mathrm{x}) = 0 \quad , \quad curl\boldsymbol{J} \equiv \gamma\, curl\boldsymbol{E}(\boldsymbol{x}) = 0 \tag{1.59}$$

inside each conductor. It follows that if $\boldsymbol{H}$ is twice differentiable

$$curlcurl\boldsymbol{H} \equiv 0$$

and that (see Proposition 1.3.1)

$$\gamma\rho(\boldsymbol{x}) \equiv 0$$

Thus all steady electric charges reside on the surface of the conductor, in the shape of surface charges with density σ. As $\partial\sigma/\partial t \equiv 0$, eq. (1.41) implies that at the interface of a conductor with a dielectric the normal component of $\boldsymbol{J}$ vanishes

$$\boldsymbol{J} \cdot \boldsymbol{n} \equiv \gamma\boldsymbol{E} \cdot \boldsymbol{n} = 0 \tag{1.60}$$

That $\boldsymbol{J}$ is solenoidal also follows from the continuity equation (M5) of §1.3.1.

In the presence of N adjoining different media Ω_j, dielectric or conducting, the equations (1.57) and (1.58) hold at the interior of every Ω_j and the equations (R1)-(R5) and (1.60) hold at the boundary surfaces between different media. We will assume that $\boldsymbol{E}(\boldsymbol{x})$,$\boldsymbol{H}(\boldsymbol{x})$,$\boldsymbol{J}(\boldsymbol{x})$ are bounded and continuous in all the Ω_j, $j = 1, ..., N$, together with all their first partial derivatives. These regularity conditions exclude the presence of point charges and electric or magnetic dipoles. Since the energy of the electromagnetic field

$$\mathcal{E}_\Omega = \int_\Omega (\frac{1}{2}\epsilon|\boldsymbol{E}(\boldsymbol{x})|^2 + \frac{1}{2}\mu|\boldsymbol{H}(\boldsymbol{x})|^2)\, \boldsymbol{dx}$$

is time-independent, if $\Omega = \underset{j}{\cup}\Omega_j$ is bounded the energy balance equation (1.51) reduces to

$$\int_\Omega \gamma|\boldsymbol{E}(\boldsymbol{x})|^2\, \boldsymbol{dx} = -\int_{\partial\Omega} \boldsymbol{E} \wedge \boldsymbol{H} \cdot \boldsymbol{n}\, dS \tag{1.61}$$

Here $\boldsymbol{n}$ is the outer normal to $\partial\Omega$ and $\epsilon \geq \epsilon_o$, $\mu > 0$ and $\gamma \geq 0$ are piecewise constant in Ω.

If $\Omega = \underset{j}{\cup}\Omega_j$ is unbounded, eq. (1.61) holds if $\boldsymbol{E}(\boldsymbol{x})$ and $\boldsymbol{H}(\boldsymbol{x})$ satisfy the condition at infinity (1.54), so that the power flux at infinity vanishes.In particular if $\underset{j}{\cup}\Omega_j = \mathbb{R}^3$ the stationary energy balance equation becomes simply

$$\int_{\mathbb{R}^3} \gamma|\boldsymbol{E}(\boldsymbol{x})|^2\, \boldsymbol{dx} = \boldsymbol{0} \tag{1.62}$$

and says that no steady electric field exists in any conductor of the family $\{\Omega_k\}$. Therefore the electric current $\boldsymbol{J}$ is zero everywhere, and across the conductors surface (where $\boldsymbol{E}\wedge\boldsymbol{n}=\boldsymbol{0}$) the power flux $\boldsymbol{S}\cdot\boldsymbol{n}$ vanishes.

We remark that eqs. (1.61) and (1.62) can be obtained directly from the two stationary equations

$$curl\,\boldsymbol{E}(\boldsymbol{x}) = 0 \quad , \quad curl\,\boldsymbol{H}(\boldsymbol{x}) = \gamma\boldsymbol{E}(\boldsymbol{x})$$

and hence they remain valid for magnetic bodies, whose constituitive relation $\boldsymbol{B}=\boldsymbol{B}(\mathbf{H})$ is nonlinear and possibly many-valued (see §1.3.1).

An alternative (equivalent) form of the steady energy balance equation based on the representation of the electric field

$$\boldsymbol{E}(\boldsymbol{x}) = -grad\, u(\boldsymbol{x})$$

in terms of an electric potential $u(\mathbf{x})$ (possibly many-valued) follows from the fact that $\boldsymbol{J}$ is solenoidal, so that

$$\int_\Omega \boldsymbol{J}\cdot\boldsymbol{E}\; dV = -\int_\Omega \mathbf{J}\cdot grad\, u\, dV = -\int_\Omega div(\boldsymbol{J}u)\, dV = \int_{\partial\Omega} u\boldsymbol{J}\cdot\boldsymbol{n}$$

We thus obtain the form of the energy equation

$$\int_\Omega \gamma|\boldsymbol{E}(\boldsymbol{x})|^2\, \boldsymbol{dx} = -\int_{\partial\Omega} u\boldsymbol{J}\cdot\boldsymbol{n}\, dS \tag{1.63}$$

perfectly equivalent to (1.61) (Exercise 15).

1.6.1 Mathematical analysis.

The stationary electric and magnetic fields in eqs. (1.57)−(1.60) are coupled solely via the current $\boldsymbol{J}(\boldsymbol{x}) \equiv \gamma\boldsymbol{E}(\boldsymbol{x})$, and so if $\boldsymbol{J}\equiv\boldsymbol{0}$ everywhere, $\boldsymbol{E}$ and $\boldsymbol{H}$ become uncoupled. In order to proceed we need the following Lemmas.

Lemma 1.6.1 *Suppose Ω is either*

(i) a surfacewise simply connected (s.s.c.) normal domain in $\mathbb{R}^3$, or

(ii) $\Omega = \mathbb{R}^3$.

Let $\boldsymbol{v}(\boldsymbol{x})$ be a bounded $C^o(\overline{\Omega}) \cap C^1(\Omega)$ vector field satisfying

$$curl\, \boldsymbol{v}(\boldsymbol{x}) = div\, \boldsymbol{v}(\boldsymbol{x}) = 0$$

in Ω, and the boundary condition

(i) $\boldsymbol{v} \wedge \boldsymbol{n} \equiv \mathbf{0}$ on $\partial\Omega$ in the first case, or

(ii) the asymptotic condition (1.54) at infinity

$$\boldsymbol{v}(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-2}) \qquad as\ |\boldsymbol{x}| \to \infty$$

if $\Omega = \mathbb{R}^3$.

Then $\boldsymbol{v}(\boldsymbol{x}) \equiv \mathbf{0}$ in Ω.

Proof.(i) Since $div\ \boldsymbol{v} = \mathbf{0}$ in Ω s.s.c. simply connected $\Rightarrow \boldsymbol{v} = curl\, \boldsymbol{V}$ with $\boldsymbol{V}$ regular in all of Ω. By applying the identity (1.48) to the vector fields $\boldsymbol{v}$, $\boldsymbol{V}$ we have

$$\boldsymbol{v} \cdot curl\ \boldsymbol{V} \ \equiv \ - div(\boldsymbol{v} \wedge \boldsymbol{V})$$

as $\boldsymbol{v}$ is irrotational. A straightforward application of the divergence theorem (DT) of §1.1 yields then

$$\int_\Omega |\boldsymbol{v}|^2 dV = \int_\Omega \boldsymbol{v} \cdot curl\, \boldsymbol{V}\, dV = -\int_\Omega div(\boldsymbol{v} \wedge \boldsymbol{V}) dV = -\int_{\partial\Omega} \boldsymbol{v} \wedge \boldsymbol{V} \cdot \boldsymbol{n}\, dS$$

and the last integral vanishes, as $\boldsymbol{v} \wedge \boldsymbol{V} \cdot \boldsymbol{n} \equiv \boldsymbol{n} \wedge \boldsymbol{v} \cdot \mathbf{V} = 0$. Thus the $L^2(\Omega)-$norm of $\boldsymbol{v}$ is zero and, since $\boldsymbol{v}$ is continuous, this implies $\boldsymbol{v}(\boldsymbol{x}) \equiv \mathbf{0}$ in Ω.

(ii), Since $curl\, \boldsymbol{v} = \mathbf{0}$ in $\mathbb{R}^3 \Rightarrow \boldsymbol{v} = -grad\, u$ with u regular (one-valued) in $\mathbb{R}^3$. As $\boldsymbol{v}$ is also solenoidal u is harmonic, i.e.

$$\Delta_3 u = 0 \qquad \text{in } \mathbb{R}^3$$

Because of the asymptotic condition (1.54) for $\boldsymbol{v}$, u is also regular at infinity [2], i.e.

$$u(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-1}) \qquad \text{as } |\boldsymbol{x}| \to \infty \tag{1.64}$$

Since $div\ \boldsymbol{v} = 0$ we have $\boldsymbol{v}\cdot grad\, u \equiv div(u\boldsymbol{v})$, and

$$\int_{\mathbb{R}^3} |\boldsymbol{v}|^2 d\boldsymbol{x} = -\lim_{R\to\infty}\int_{|\boldsymbol{x}|<R} \boldsymbol{v}\cdot grad\, u\, d\boldsymbol{x} = -\lim_{R\to\infty}\int_{|\boldsymbol{x}|=R} u\boldsymbol{v}\cdot \boldsymbol{n}\, dS_{\boldsymbol{x}} = 0$$

as $u\boldsymbol{v} = O(|\boldsymbol{x}|^{-3})$ as $|\boldsymbol{x}| \to \infty$. Thus $\boldsymbol{v}(\boldsymbol{x}) \equiv \mathbf{0}$.

Lemma 1.6.2 *Suppose* Ω *is either*

(i) *a contourwise simply connected (c.s.c.) normal domain in* $\mathbb{R}^3$, *or*

(ii) *the c.s.c. domain exterior to a closed connected surface* S *in* $\mathbb{R}^3$, *or*

(iii) *a simply connected normal domain in* $\mathbb{R}^3$. *Let* $\boldsymbol{v}(\boldsymbol{x})$ *be a bounded* $C^o(\overline{\Omega}) \cap C^1(\Omega)$ *vector field satisfying*

$$curl\boldsymbol{v}(\boldsymbol{x}) = div\boldsymbol{v}(\boldsymbol{x}) = 0 \quad in\ \Omega$$

and the boundary condition

(i) $\boldsymbol{v}\cdot\boldsymbol{n} = \mathbf{0}$ *on* $\partial\Omega$ *in the first case ,*

(ii) $\boldsymbol{v}\cdot\boldsymbol{n} = \mathbf{0}$ *on* S *and the asymptotic condition* $\boldsymbol{v}(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-2})$ *as* $|\boldsymbol{x}| \to \infty$, *in the second case,*

(iii) $\boldsymbol{v}\cdot\boldsymbol{n} = \mathbf{0}$ *on* S_1 *and* $\boldsymbol{v}\wedge\boldsymbol{n} \equiv \mathbf{0}$ *on* S_2 , $S_1 \cup S_2 = \partial\Omega, S_1 \cap S_2 = \emptyset$, *in the third case.*

Then $\boldsymbol{v}(\boldsymbol{x}) \equiv \mathbf{0}$ *in* Ω.

Proof. As $curl\ \boldsymbol{v} = \mathbf{0}$ in Ω c.s.c. $\Rightarrow \boldsymbol{v} = -grad\, u$ with u regular (one-valued) in Ω. As $\boldsymbol{v}$ is also solenoidal $u(\boldsymbol{x})$ is harmonic in Ω and satisfies the Neumann boundary condition $\partial u/\partial n = 0$ on $\partial\Omega$.

(i) Since $\mathbf{v}\cdot grad\, u \equiv div(u\boldsymbol{v})$ we obtain

$$\int_{\Omega} |\boldsymbol{v}|^2 dV = -\int_{\Omega} \boldsymbol{v}\cdot grad\, u\, d\boldsymbol{x} = -\int_{\partial\Omega} u\boldsymbol{v}\cdot\boldsymbol{n}\, dS \equiv \int_{\partial\Omega} u\frac{\partial u}{\partial n}\, dS = 0$$

whence $\mathbf{v}(\mathbf{x}) \equiv \mathbf{0}$ in Ω.

(ii) u satisfies the condition $u(\mathbf{x}) = O(|\mathbf{x}|^{-1})$ as $|\mathbf{x}| \to \infty$. Let $\Omega_R = \{\mathbf{x} \in \Omega : |\mathbf{x}| < R\}$. Proceeding as in part (ii) of Lemma 1.6.1 we then obtain

$$\int_{\Omega} |\boldsymbol{v}|^2 d\boldsymbol{x} = -\lim_{R\to\infty}\int_{\Omega_R} \boldsymbol{v}\cdot grad\, u\, d\boldsymbol{x} = -\lim_{R\to\infty}\int_{\partial\Omega_R} u\boldsymbol{v}\cdot\boldsymbol{n}\, dS_x = 0$$

as $\partial\Omega_R = \partial\Omega \cup \{|\boldsymbol{x}| = R\}$, $\boldsymbol{v}\cdot\boldsymbol{n}= 0$ on $\partial\Omega$ and $u\boldsymbol{v}= O(|\boldsymbol{x}|^{-3})$ as $|\boldsymbol{x}| \to \infty$. Therefore $\boldsymbol{v}(\boldsymbol{x}) \equiv \boldsymbol{0}$ in Ω.

(*iii*) In this case Ω is both c.s.c. and s.s.c., $\partial\Omega$ is a connected surface and either $\boldsymbol{v}\cdot\boldsymbol{n}= 0$ or (if $\boldsymbol{v}\wedge\boldsymbol{n}\equiv \boldsymbol{0}$) $u =$M constant on $\partial\Omega$. We have just shown that

$$\int_\Omega |\boldsymbol{v}|^2 dV = -\int_{\partial\Omega} u\boldsymbol{v}\cdot\boldsymbol{n}\, dS$$

On the other hand by the divergence theorem we obtain

$$0 = \int_\Omega div\,\boldsymbol{v}\, dV = \int_{\partial\Omega} \boldsymbol{v}\cdot\boldsymbol{n}\, dS = \int_{S_2} \boldsymbol{v}\cdot\boldsymbol{n}\, dS$$

Therefore, taking into account the boundary condition for $\boldsymbol{v}$ and u yields

$$\int_{\partial\Omega} u\boldsymbol{v}\cdot\boldsymbol{n}\, dS = \mathrm{M}\int_{S_2} \boldsymbol{v}\cdot\boldsymbol{n}\, dS = 0$$

and so $\boldsymbol{v}(\boldsymbol{x}) \equiv \boldsymbol{0}$ in Ω.

If Ω is not simply connected these results may not be true, as the following counterexamples show.

Counterexample 1. If Ω is contourwise multiply connected, for example if Ω is bounded by a torus or a toroidal surface, there exist nonzero vector fields $\boldsymbol{J}$, irrotational and solenoidal in Ω, and with $\boldsymbol{J}\cdot\boldsymbol{n}\equiv \boldsymbol{0}$ on $\partial\Omega$. They are usually called Neumann vector fields and can be written as gradients $\boldsymbol{J}= -grad\, u$ of a harmonic many-valued scalar potential u in Ω, so that the boundary condition $\partial u/\partial n = 0$ on $\partial\Omega$ does not imply that $u =$constant in Ω [37].

Counterexample 2. Similarly, if Ω is surfacewise multiply connected, for example if Ω is the domain bounded by two concentric spheres (spherical condenser), there exist irrotational and solenoidal vector fields $\boldsymbol{E}$ that can not be written as the *curl* of a one-valued vector potential $\boldsymbol{V}$ globally in Ω. Thus $\boldsymbol{E}= -grad\, u$, where the scalar potential u is harmonic and single-valued in Ω but may assume different boundary values on the different connected components of $\partial\Omega$, so that the boundary condition $\boldsymbol{E}\wedge\boldsymbol{n}\equiv \boldsymbol{0}$ on $\partial\Omega$ does not imply that $\boldsymbol{E}\equiv \boldsymbol{0}$ in Ω (Exercise 16).

Lemma 1.6.3 *Let Ω be a s.s.c. normal domain in $\mathbb{R}^3$ and $\Omega' := \mathbb{R}^3\backslash\overline{\Omega}$, and let $v(\boldsymbol{x})$ be a function harmonic and biregular in Ω and in Ω', and regular at* *infinity* [10]:

$$O\left(|\boldsymbol{x}|^{-1}\right) \quad as\ |\boldsymbol{x}| \to \infty$$

Then if $\partial v/\partial n$ is continuous across $\partial\Omega$ and the jump of v is constant on $\partial\Omega$ $\Rightarrow$ $grad\, v \equiv 0$ in Ω and in Ω'.

Proof. Any function $v(\boldsymbol{x})$ biregular in Ω and in Ω', regular at infinity, and whose Laplacian $\Delta_3 v$ has compact support $K \subset \Omega$, can be represented via the Green identity

$$\text{(GI)} \qquad v(\boldsymbol{x}) = \frac{1}{4\pi}\int_{\partial\Omega}\Big\{[v(\boldsymbol{y})]_{\partial\Omega}\,\frac{\partial}{\partial n_y}\frac{1}{r} - [\frac{\partial v}{\partial n_y}]_{\partial\Omega}\,\frac{1}{r}\Big\}dS_y - \frac{1}{4\pi}\int_{\Omega}\frac{1}{r}\Delta_3 v(\boldsymbol{y})\,\boldsymbol{dy}$$

for $\boldsymbol{x} \in \Omega\cup\Omega'$, and

$$v_-(\boldsymbol{x}) + v_+(\boldsymbol{x}) = \frac{1}{2\pi}\int_{\partial\Omega}\Big\{[v(\boldsymbol{y})]_{\partial\Omega}\,\frac{\partial}{\partial n_y}\frac{1}{r} - [\frac{\partial v}{\partial n_y}]_{\partial\Omega}\,\frac{1}{r}\Big\}dS_y - \frac{1}{2\pi}\int_{\Omega}\frac{1}{r}\Delta_3 v(\boldsymbol{y})\,\boldsymbol{dy}$$

for $\boldsymbol{x} \in \partial\Omega$ [37]. We denote $r = |\boldsymbol{x} - \boldsymbol{y}|$, $[v]_{\partial\Omega} = v_+ - v_-$, $\boldsymbol{n}$ the outer normal to $\partial\Omega$, a connected surface. Since by assumption $\Delta_3 v \equiv 0$, $[\frac{\partial v}{\partial n_y}]_{\partial\Omega} \equiv 0$ and $[v(\boldsymbol{y})]_{\partial\Omega} \equiv \mathrm{M}$, a constant, (GI) reduces to

$$v(\boldsymbol{x}) = \frac{\mathrm{M}}{4\pi}\int_{\partial\Omega}\frac{\partial}{\partial n_y}\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|}\; dS_y$$

and by the Gauss solid angle formula [2] $v(\boldsymbol{x})$ is piecewise constant:

$$v(\boldsymbol{x}) = \begin{cases} -\mathrm{M} & \boldsymbol{x} \in \Omega \\ 0 & \boldsymbol{x} \in \Omega' \end{cases}$$

Therefore $grad\, v \equiv 0$ in Ω and in Ω', as asserted.

Corollary 1.6.4 *Let Ω be a s.s.c. normal domain in $\mathbb{R}^3$, $\Omega' := \mathbb{R}^3 \backslash \overline{\Omega}$, and let $\boldsymbol{v}(\boldsymbol{x})$ be a vector function, continuous in $\mathbb{R}^3$ with continuous bounded derivatives in $\Omega\cup\Omega'$, and satisfying the asymptotic condition (1.54) at infinity. Then if $curl\ \boldsymbol{v} = div\ \boldsymbol{v} = 0$ in $\Omega\cup\Omega'$ $\Rightarrow$ $\boldsymbol{v} \equiv \boldsymbol{0}$ everywhere.*

[10] this condition is the counterpart of the asymptotic condition (1.54) for the vector field $-grad\, v$

Proof. We have $\boldsymbol{v}= -grad\, v$ where the potential v satisfies all the assumptions of Lemma 1.6.3. In particular v is one-valued even if Ω is c.m.c., and $[v(\boldsymbol{y})]_{\partial\Omega}$ =constant as a consequence of eq. (1.43) and the fact that $\partial\Omega$ is connected. Hence v$\equiv \mathbf{0}$ in $\Omega \cup \Omega'$ and, by continuity, also on $\partial\Omega$.

We have seen that the stationary current density $\boldsymbol{J}$ in a conductor C satisfies [11]

$$div\boldsymbol{J} = curl\boldsymbol{J} = 0 \quad \text{in } C \quad , \qquad \boldsymbol{J}\cdot\boldsymbol{n} = 0 \ \text{ on } \partial C$$

(see eqs. (1.59) and (1.60)). The two following theorems imply that, if the power flux at infinity vanishes, energy dissipation makes a non-vanishing steady current and electric field impossible in any conductor.

Theorem 1.6.5 *Let C be a conductor surrounded by an unbounded dielectric D with permittivity ϵ extending to infinity in $\mathbb{R}^3$. Suppose $\boldsymbol{E}(\boldsymbol{x})$, $\boldsymbol{H}(\boldsymbol{x})$ satisfy eqs. (1.57)–(1.60) and, if C is c.m.c., also the asymptotic condition at infinity (1.54). Then*

$$\boldsymbol{E}(\boldsymbol{x}) \equiv \boldsymbol{J}(\boldsymbol{x}) \equiv \ \boldsymbol{D}(\boldsymbol{x}) \equiv \rho(\boldsymbol{x}) \equiv 0 \qquad in\ \overline{C}$$

On the conductor surface ∂C we have

$$\boldsymbol{E}\wedge\boldsymbol{n} = \mathbf{0} \tag{1.65}$$

and also [12]

$$\boldsymbol{E}\cdot\boldsymbol{n} = \frac{\sigma}{\epsilon} \tag{1.66}$$

where σ is the surface charge density over the conductor surface ∂C, and the normal $\boldsymbol{n}$ to $\partial C = \partial D$ is oriented towards D.

Proof. If C is c.s.c., Lemma 1.6.2 and eq. (1.60) imply that $\boldsymbol{J}\equiv \mathbf{0}$ in C; if C is c.m.c., the same conclusion follows from eq. (1.62). Thus inside C and, by continuity, up to the inner conductor boundary we have $\boldsymbol{E}= \gamma^{-1}\boldsymbol{J}\equiv \mathbf{0}$ by (C1), $\boldsymbol{D}= \epsilon\boldsymbol{E}\equiv \mathbf{0}$ by (C2), and $\rho = div\boldsymbol{D}\equiv 0$ by (M3). Since $\boldsymbol{E}\wedge\boldsymbol{n}$ is continuous across ∂C by virtue of (R5), eq. (1.65) follows. Finally eq. (1.66) is a consequence of the matching relation (R1), $[\boldsymbol{D}]_{\partial C}\cdot\boldsymbol{n}= \sigma$, the constituitive equation (C2), $\boldsymbol{D}= \epsilon\boldsymbol{E}$, and the fact that $\boldsymbol{E}\equiv \mathbf{0}$ in $\overline{C}$.

[11] for brevity we denote by the same letter the material and the domain it occupies

[12] the normal trace $\mathbf{E}\cdot\mathbf{n}$ is defined as the limit as the point approaches ∂C coming from the dielectric (in the conductor $\mathbf{E}= \mathbf{0}$)

Theorem 1.6.6 *Suppose that the power flux at infinity vanishes so that the stationary energy balance equation (1.62) holds in* $\mathbb{R}^3$

$$\int_{\mathbb{R}^3} \gamma |\boldsymbol{E}(\boldsymbol{x})|^2 \, \boldsymbol{dx} = 0$$

and suppose that $\mathbb{R}^3$ *is the union of any number of disjoint material media, conducting and/or dielectric. Then*

$$\boldsymbol{J}(\boldsymbol{x}) \equiv \boldsymbol{0} \quad \textit{in } \mathbb{R}^3 \,, \quad \boldsymbol{E}(\boldsymbol{x}) \equiv \boldsymbol{0} \quad \textit{in all the conductors}$$

and the stationary power flux across any closed surface contained in $\mathbb{R}^3$ *is zero*

$$\oint_S \boldsymbol{S} \cdot \boldsymbol{n} \, dS = 0 \qquad\qquad \forall S,\, \partial S = \emptyset$$

so that

$$\boldsymbol{E}(\boldsymbol{x}) \wedge \boldsymbol{H}(\boldsymbol{x}) \equiv \boldsymbol{0} \quad \textit{in } \mathbb{R}^3$$

Proof. Eq. (1.62) implies that $\boldsymbol{E} \equiv \boldsymbol{0}$ in all the conductors, where $\gamma > 0$, so that $\boldsymbol{J} \equiv \boldsymbol{0}$ everywhere. By Proposition 1.5.1, $\boldsymbol{S} \cdot \boldsymbol{n}$ is continuous across any surface in $\mathbb{R}^3$. The energy balance equation (1.6.1) applied to the domain Ω bounded by S becomes then

$$\oint_S \boldsymbol{E} \wedge \boldsymbol{H} \cdot \boldsymbol{n} \, dS = 0$$

As S is arbitrary, it follows that $\boldsymbol{E} \wedge \boldsymbol{H} = \boldsymbol{0}$ everywhere.

Thus under the stated assumptions a conductor cannot support a steady electric current and can carry only a surface charge σ distributed over the separation surface with a dielectric according to eq. (1.66).

The condition $\boldsymbol{E}(\boldsymbol{x}) \wedge \boldsymbol{H}(\boldsymbol{x}) \equiv \boldsymbol{0}$ in Theorem 1.6.6 is satisfied if either $\boldsymbol{E}$ or $\boldsymbol{H}$ are identically zero. We give here an example where $\boldsymbol{H} \equiv \boldsymbol{0}$.

Example. Consider a conducting ball $C := \{\boldsymbol{x} \in \mathbb{R}^3 \colon |\boldsymbol{x}| < R\}$ surrounded by a dielectric $D := \mathbb{R}^3 \setminus \overline{C}$ with permittivity ϵ. Let

$$\rho(\boldsymbol{x}) \equiv 0 \ \text{ for } |\boldsymbol{x}| > R \ , \qquad \sigma(\boldsymbol{x}) = \sigma_o \ \text{ for } |\boldsymbol{x}| = R$$

The unique fields satisfying the asymptotic condition (1.54), so that eq. (1.62) holds in $\mathbb{R}^3$, are

$$\boldsymbol{H} \equiv \boldsymbol{0} \ \ , \boldsymbol{E} = \begin{cases} \boldsymbol{0} & \text{in } C : |\boldsymbol{x}| \leq R \\ -grad \frac{R^2 \sigma_o}{\epsilon |\boldsymbol{x}|} & \text{in } D : |\boldsymbol{x}| > R \end{cases}$$

(Exercise 17).

The next propositions concern situations where either $\boldsymbol{E}\equiv \mathbf{0}$ or $\boldsymbol{H}\equiv \mathbf{0}$ or both.

Proposition 1.6.7 *Suppose that the entire space $\mathbb{R}^3$ is filled by a single homogeneous medium with $\gamma \geq 0$, and that $\boldsymbol{E}(\boldsymbol{x})$ and $\boldsymbol{H}(\boldsymbol{x})$ satisfy the asymptotic condition (1.54) at infinity. Then*

(i) $\boldsymbol{H}\equiv \mathbf{0}$ *in* $\mathbb{R}^3$

(ii) If $\rho(x) \equiv 0 \Rightarrow \boldsymbol{E}\equiv \mathbf{0}$ *in* $\mathbb{R}^3$.

Proof. (i) Since $\mathbf{J}\equiv \mathbf{0}$ in $\mathbb{R}^3$ by assumption (if $\gamma = 0$) or from eq. (1.62) (if $\gamma = 0$), $\boldsymbol{H}$ satisfies the equation $curl\,\boldsymbol{H}(\boldsymbol{x}) = div\boldsymbol{H}(\boldsymbol{x}) = 0$ in $\mathbb{R}^3$ from eq. (1.58), and the condition (1.54) by assumption. Hence $\boldsymbol{H}\equiv \mathbf{0}$ in $\mathbb{R}^3$ by Lemma 1.6.1(ii).

(ii) If $\gamma > 0$ we have $\boldsymbol{E}\equiv \mathbf{0}$ in $\mathbb{R}^3$ from eq. (1.62). If $\gamma = 0$ then $curl\,\boldsymbol{E}(\boldsymbol{x}) = div\ \boldsymbol{E}(\boldsymbol{x}) = 0$ in $\mathbb{R}^3$ from eq. (1.57) with $\rho \equiv 0$ and, as $\boldsymbol{E}$ satisfies (1.54) by assumption, Lemma 1.6.1(ii) implies that $\boldsymbol{E}\equiv 0$.

Proposition 1.6.8 *Suppose that Ω is a s.s.c. normal domain in $\mathbb{R}^3$ and that Ω' is the exterior domain $\Omega' = \mathbb{R}^3\backslash\overline{\Omega}$. Suppose that Ω, Ω' are filled by two homogeneous media one of which is conducting. Let $\rho(\boldsymbol{x}) \equiv 0$ and* **E, H** *satisfy the asymptotic condition (1.54) in Ω'. Then*:

(i) Ω *a conductor,* Ω' *a dielectric and* $\sigma(\boldsymbol{x}) \equiv 0$ *on* $\partial\Omega \Rightarrow \boldsymbol{E}\equiv \mathbf{0}$ *in* $\mathbb{R}^3$

(ii) Ω *a dielectric,* Ω' *a conductor* $\Rightarrow \sigma(\boldsymbol{x}) \equiv 0$ *on* $\partial\Omega$ *and* $\boldsymbol{E}\equiv \mathbf{0}$ *in* $\mathbb{R}^3$.

Proof. By assumption $\partial\Omega$ is connected, and Ω' is s.s.c. like Ω.

(i) As $\rho(\boldsymbol{x}) \equiv \sigma(\boldsymbol{x}) \equiv 0$ we have that $\boldsymbol{E}(\boldsymbol{x})$ satisfies all assumptions of Corollary 1.6.4. Hence $\boldsymbol{E}\equiv \mathbf{0}$.

(ii) $\boldsymbol{E}\equiv \mathbf{0}$ in Ω' from eq. (1.63) applied to $\mathbb{R}^3$. By continuity $\boldsymbol{E}\wedge\boldsymbol{n}= \mathbf{0}$ on $\partial\Omega$, and as $\rho(\boldsymbol{x}) \equiv 0$ we have $curl\,\boldsymbol{E}(\boldsymbol{x}) = div\ \boldsymbol{E}(\boldsymbol{x}) = 0$ in Ω by (1.57). As Ω is s.s.c., Lemma 1.6.1 implies that $\boldsymbol{E}\equiv \mathbf{0}$ in Ω and $\boldsymbol{E}\cdot\boldsymbol{n}= \mathbf{0}$ on $\partial\Omega$, hence $\sigma \equiv 0$ from eq. (1.66).

Proposition 1.6.9 *Suppose that* Ω *is a s.s.c. normal domain in* $\mathbb{R}^3$ *filled by a homogeneous conductor or dielectric and surrounded by a homogeneous dielectric in* $\Omega' := \mathbb{R}^3 \backslash \overline{\Omega}$ *. Suppose that* $\boldsymbol{J} \equiv \mathbf{0}$ *everywhere, or equivalently, that* $\boldsymbol{E}, \boldsymbol{H}$ *satisfy (1.54) at infinity.*

Then $\mu = \mu_o$ *everywhere* $\Rightarrow \boldsymbol{H} \equiv \mathbf{0}$ *in* $\mathbb{R}^3$*.*

Proof. We have $\boldsymbol{J} \equiv \mathbf{0}$ everywhere by assumption or from eq. (1.62) applied to $\mathbb{R}^3$ (if Ω is a conductor). Therefore $curl\, \boldsymbol{H}(\boldsymbol{x}) = div \boldsymbol{H}(\boldsymbol{x}) = 0$ in $\Omega \cup \Omega'$ by (1.55). By the assumption $\mu = \mu_o$ and by (R2) and (R4), $\boldsymbol{H}$ is continuous in $\mathbb{R}^3$. Corollary 1.6.4 implies then that $\boldsymbol{H} \equiv \mathbf{0}$ in $\mathbb{R}^3$.

Note that Proposition 1.6.9 does not apply to permanent magnets or superconductors, as it makes use of the linear constitutive relation $\boldsymbol{B} = \mu_o \boldsymbol{H}$.

The conclusions of Theorem 1.6.5, Theorem 1.6.6, Proposition 1.6.8 and Proposition 1.6.9 can be summarized by the following [13]

Corollary 1.6.10 *If the power flux at infinity vanishes* $\boldsymbol{J}(\boldsymbol{x}) \equiv \mathbf{0}$ *everywhere,* $\boldsymbol{E}(\boldsymbol{x}) \equiv \mathbf{0}$ *in the absence of all charges, and* $\boldsymbol{H}(\boldsymbol{x}) \equiv \mathbf{0}$ *in the absence of magnetic materials.*

This result motivates the definition of two important subclasses of stationary fields, known as Electrostatics and Magnetostatics. They are characterized by the fact that the electric current and the Poynting vector vanish, so that the electric and magnetic fields are uncoupled. Again, we restrict our attention to one or several homogeneous media, so that $\epsilon \geq \epsilon_o$, $\mu > 0$ and $\gamma \geq 0$ are piecewise constant in $\mathbb{R}^3$.

1.6.2 Electrostatics.

We assume that $\boldsymbol{B} \equiv \boldsymbol{H} \equiv \boldsymbol{J} \equiv \mathbf{0}$. Then $\boldsymbol{S} \equiv \mathbf{0}$ and inside any homogeneous medium with permittivity ϵ, the electrostatic field $\boldsymbol{E}(\boldsymbol{x})$ satisfies

$$curl\, \boldsymbol{E}(\boldsymbol{x}) = \mathbf{0} \quad , \qquad div \boldsymbol{E}(\boldsymbol{x}) = \rho(\boldsymbol{x})/\epsilon$$

[13] we recall that our assumptions exclude the presence of singularities like point charges and electric or magnetic dipoles

It follows that $\boldsymbol{E} = -grad\, u(\boldsymbol{x})$ where the electrostatic potential $u(\boldsymbol{x})$, possibly many-valued, satisfies the Poisson equation

$$\Delta_3\, u = -\frac{\rho(\boldsymbol{x})}{\epsilon} \tag{P}$$

in a dielectric D, and the Laplace equation

$$\Delta_3\, u = 0 \tag{L}$$

in a conductor C, where $\rho(\boldsymbol{x}) \equiv 0$. In fact, at the interior of any conductor C we have $\boldsymbol{E} = \gamma^{-1}\boldsymbol{J} \equiv \mathbf{0}$, $\boldsymbol{D} \equiv \mathbf{0}$ (cfr. Theorem 1.6.5), hence $u(\boldsymbol{x})$ is constant:

$$u(\boldsymbol{x}) = u_- \equiv \text{constant} \qquad for\, \boldsymbol{x} \in C$$

and is thus trivially a harmonic function. By continuity the inner limit of $u(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches ∂C on the conductor side, say the $-$ side, is given by the same constant u_-

$$\lim_{\substack{\boldsymbol{x}\to\partial C\\ \boldsymbol{x}\in C}} u(\boldsymbol{x}) = u_- \tag{1.67}$$

(we recall that C is connected by definition). On the other hand we have seen before (cfr. eq. (1.43)) that the potential jump $[u]_{\partial C}$ is constant on each connected component of ∂C. Therefore the outer limit of u on ∂C

$$u_+ := \lim_{\substack{\boldsymbol{x}\to\partial C\\ \boldsymbol{x}\in C'}} u(\boldsymbol{x}) \qquad\qquad (C' = \mathbb{R}^3 \backslash \overline{C})$$

satisfies

$$u_+ = u_C \quad , \quad u_C := u_- + [u]_{\partial C} \tag{1.68}$$

and is constant on each connected component of ∂C. (The value of u_C may vary from one connected component of ∂C to the other.) From eqs. (1.65) and (1.66), $\mathbf{E}\wedge\mathbf{n} = \mathbf{0}$ over the conductor surface ∂C and if the conductor is surrounded by a dielectric, $\boldsymbol{E}\cdot\boldsymbol{n} = \sigma/\epsilon$ over the outer side of the conductor surface, so the lines of force of $\boldsymbol{E}$ and $\boldsymbol{D}$ in the dielectric are orthogonal to the conductor surface. It follows that $\boldsymbol{S}\cdot\boldsymbol{n} = 0$ and the power flux across the conductor surface vanishes.

To summarize: in electrostatics a conductor C may carry only a surface charge, distributed over the boundary between C and a dielectric D, with density σ given by

$$\sigma(\boldsymbol{y}) = -\epsilon \frac{\partial u}{\partial n_y} := -\epsilon \lim_{\substack{\boldsymbol{x}\to\boldsymbol{y} \\ \boldsymbol{x}\in D}} \boldsymbol{n}(\mathbf{y}) \cdot grad_x u(\boldsymbol{x}) \qquad (\boldsymbol{y} \in \partial D) \tag{1.69}$$

where $\boldsymbol{n}$ is the unit outer normal to ∂C, oriented towards D, and ϵ is the permittivity of the dielectric. Inside a conductor the electric field vanishes and the electrostatic potential u has a constant value u_-. On the outer side of the surface of C the electrostatic potential satisfies the boundary condition $u|_{\partial C} = u_C$, interpreted in the sense of the limit from the exterior of C, or from the interior of the dielectric:

$$\lim_{\substack{\mathbf{x}\to\partial C \\ \boldsymbol{x}\in D}} u(\boldsymbol{x}) = u_C$$

The outer limit u_C is constant over each connected component of ∂C and is related to the inner value u_- of u and to the possible jump $[u]_{\partial C}$ of u across ∂C by the second equation (1.68). Thus, if $[u]_{\partial C} = 0$ over the entire boundary ∂C, the outer and inner limits u_C and u_- coincide and the constant u_C is the same for all boundary ∂C.

More generally, if N conductors C_j are immersed in a dielectric D, the electrostatic potential u is constant inside each C_j

$$u\,(\boldsymbol{x}) = u_{_j} = \text{ constant } \ for\ \boldsymbol{x} \in C_j$$

and satisfies the N boundary conditions

$$\lim_{\substack{\boldsymbol{x}\to\partial C_j \\ \boldsymbol{x}\in D}} u(\boldsymbol{x}) = u_{C_j} \equiv u_{_j} + [u]_{\partial C_j} \qquad (j = 1, ..., N) \tag{1.70}$$

where the jump $[u]_{\partial C_j}$ is constant along each connected component of ∂C_j. In other words, conductor surfaces are equipotential and u_{C_j} represents the value of u over the $j-$th conductor surface as “seen” from the dielectric surrounding it.

Since $\boldsymbol{H}\equiv\boldsymbol{S}\equiv \mathbf{0}$, the electrostatic energy in any domain Ω (which may comprise one or several or all material bodies)

$$\mathcal{E}_\Omega := \frac{1}{2}\int_\Omega \epsilon|\boldsymbol{E}(\boldsymbol{x})|^2 \boldsymbol{dx} \tag{1.71}$$

is constant in time and the stationary energy balance equation (1.61) reduces to eq. (1.62)

$$\int_\Omega \gamma |\boldsymbol{E}(\boldsymbol{x})|^2 \, \boldsymbol{dx} = \boldsymbol{0}$$

where ϵ, γ are piecewise constant in Ω with $\epsilon \geq \epsilon_o$, $\gamma \geq 0$. As we already know, this equation simply says that the stationary electrostatic field and electrostatic energy in a conductor are zero.

1.6.3 Magnetostatics.

We assume that $\boldsymbol{D}\equiv\boldsymbol{E}\equiv\boldsymbol{J}\equiv \boldsymbol{0}$. Then again $\boldsymbol{S}\equiv \boldsymbol{0}$ and inside every homogeneous medium $\boldsymbol{H}(\boldsymbol{x})$ and $\boldsymbol{B}(\boldsymbol{x})$ satisfy

$$curl\,\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{0} \qquad , \qquad div\,\boldsymbol{B}(\boldsymbol{x}) = 0$$

Thus $\boldsymbol{H}= -grad\,v(\boldsymbol{x})$ and the magnetostatic potential $v(\boldsymbol{x})$, possibly many-valued, satisfies the Laplace equation for harmonic functions

$$\text{(L)} \qquad \Delta_3\, v = 0$$

inside any material medium. The magnetostatic energy in a domain Ω is given by

$$\text{(1.72)} \qquad \mathcal{E}_\Omega = \int_\Omega \frac{1}{2}\mu|\mathbf{H}(\mathbf{x})|^2\,\mathbf{dx}$$

where μ is piecewise constant, and no energy balance equation is available.

In the case of soft iron μ is very large (see §1.8) so that $|\boldsymbol{H}| = |\boldsymbol{B}|/\mu \cong 0$ inside the body, and, since $\boldsymbol{H}\wedge\boldsymbol{n}= \boldsymbol{0}$ and $\boldsymbol{B}\cdot\boldsymbol{n}$ are continuous across the body surface, the lines of force of $\boldsymbol{H}$ and $\boldsymbol{B}$ outside the body are orthogonal to its boundary. Thus soft iron is, in a sense, the analogue of a conductor in electrostatics. (In contrast, in the case of a superconducting body $|\boldsymbol{B}| = \mu|\boldsymbol{H}| = 0$ inside the body, and the lines of force of $\boldsymbol{H}$ and $\boldsymbol{B}$ outside the body are tangent to its boundary.)

Corollary 1.6.10 above implies that, if the power flux at infinity vanishes, Magnetostatics or Electrostatics are the only possible outcome of the steady Maxwell equations (1.57), (1.58) in the absence of (ferro)magnetic materials or in the absence of all charges, respectively.

1.7 Quasi-stationary fields in conductors

Consider a homogeneous conductor C surrounded by a dielectric D. If the electric field varies "slowly" with time, the displacement current $\mathbf{J}_d$ in the conductor turns out to be negligible with respect to the conduction current $\boldsymbol{J}$

$$\left|\frac{\partial \boldsymbol{D}}{\partial t}\right| \ll |\boldsymbol{J}|$$

The quasi-stationary (or parabolic) approximation of the Maxwell equations consists in neglecting the term $\partial \boldsymbol{D}/\partial t$, that is, neglecting the displacement current. For a homogeneous non-magnetic conductor the Maxwell system becomes then

$$\mu_o \frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E}\ ,\ curl \boldsymbol{H}\ = \boldsymbol{J} \equiv \gamma\, \boldsymbol{E}\ \ ,\ \ div \boldsymbol{H} = div \boldsymbol{E} = 0 \tag{1.73}$$

and the boundary conditions at the interface $\mathbb{S}$ between C (say, $-$ side) and D ($+$ side) follow from (R2), (R4) and (R5):

$$[\boldsymbol{E}]_{\mathbb{S}} \wedge \boldsymbol{n} = \mathbf{0} \qquad , \qquad [\boldsymbol{H}]_{\mathbb{S}} = \mathbf{0} \tag{1.74}$$

The second equation (1.73) implies that $\boldsymbol{J}$ is solenoidal:

$$div \boldsymbol{J} = 0$$

as well as $\boldsymbol{E}$. Suppose that $\boldsymbol{E}$ and $\boldsymbol{H}$ are twice differentiable. Substituting the expression

$$\boldsymbol{E} = \gamma^{-1}\, curl\, \boldsymbol{H}$$

in the first eq. (1.73) yields the vector heat equation for $\boldsymbol{H}$

$$\frac{\partial \boldsymbol{H}}{\partial t} = -\frac{1}{\mu\gamma} curl\, curl\, \boldsymbol{H} \tag{1.75}$$

Using the vector identity (valid for cartesian components)

$$\Delta_3 \boldsymbol{H} \equiv grad\, div\, \boldsymbol{H}\ - curlcurl\, \boldsymbol{H} \tag{1.76}$$

and taking into account $div\, \boldsymbol{H} = 0$, eq. (1.75) takes the form of the ordinary heat equation for each cartesian component H_j of $\boldsymbol{H}$

$$\frac{\partial H_j}{\partial t} = \frac{1}{\mu\gamma} \Delta_3 H_j \qquad\qquad (j = 1, 2, 3)$$

Similarly, eliminating $\boldsymbol{H}$ from (1.73) yields the parabolic heat equation for each cartesian component of $\boldsymbol{J}$

$$\frac{\partial \boldsymbol{J}}{\partial t} = \frac{1}{\mu\gamma}\Delta_3\, \boldsymbol{J} \tag{1.77}$$

and so also of $\boldsymbol{E}=\boldsymbol{J}/\gamma$.

These equations imply that the conduction current propagates along the conductor with infinite speed in the parabolic approximation [2,31]. In spite of this apparent paradox, Maxwell's model remains valid also in the quasi-stationary approximation and yields results in full agreement with experiments. An important phenomenon which is predicted correctly by the parabolic approximation goes under the name of skin effect. This phenomenon consists in the fact that alternate currents penetrate only in a thin layer near to the boundary surface of the conductor and are damped out in the interior, the more so the higher the frequency. In order to see this, consider an unbounded,homogeneous conductor C in the half-space $z > 0$ and suppose that a time-periodic current propagates along the $x-$axis in the conductor, so that

$$\boldsymbol{J} = J(y, z, t)\boldsymbol{c}_1$$

(the scalar function J cannot depend on x because of $div\boldsymbol{J} = \partial J/\partial x \equiv 0$). Furthermore, suppose that *J(y,z,t)* is bounded and satisfies the boundary condition

$$J(y, 0, t) = Ae^{i\omega t}$$

at the interface $z = 0$ with a dielectric medium which fills the half-space $z < 0$. Since $\boldsymbol{J}\cdot\boldsymbol{n}= 0$, eq. (R3) shows that the electric surface charge density σ over the interface is necessarily zero or constant in time. The function $J(y, z, t)$ must then be determined as the bounded solution of the BVP for the heat equation in the half-space $z > 0$

$$\frac{\partial J}{\partial t} = \frac{1}{\mu\gamma}(\frac{\partial^2 J}{\partial y^2} + \frac{\partial^2 J}{\partial z^2}) \qquad z > 0,\ t \in \mathbb{R}$$
$$J(y, 0, t) = Ae^{i\omega t} \qquad t \in \mathbb{R}$$

Because the bounded solution to this problem is unique [2], $J = J(z, t)$ does not depend on y and we easily find

$$J = Ae^{-z/\delta} exp(i(\omega t - \frac{z}{\delta})) \tag{1.78}$$

where

$$\delta := \sqrt{\frac{2}{\gamma\mu_o\omega}} \tag{1.79}$$

(Exercise 18). The solution corresponding to generic boundary data $J(y,0,t) = f(t)$ periodic with period $T = 2\pi/\omega$ can be found by superposing the monochromatic solutions (1.78) via a Fourier expansion.

Eq. (1.78) shows that $\boldsymbol{J}$ is exponentially damped for increasing $z > 0$, so that $\mathbf{J} \cong \mathbf{0}$ for (say) $z \geq 10\,\delta$, and the damping increases for increasing ω. A similar *skin effect* occurs also for the electric and magnetic fields in the conductor, given by

$$\boldsymbol{E} = \frac{1}{\gamma} J(z,t)\boldsymbol{c}_1 \quad , \quad \boldsymbol{H} = \frac{e^{-i\pi/4}}{\sqrt{\omega\mu_o\gamma}} J(z,t)\boldsymbol{c}_2$$

Note that $\boldsymbol{H}$ has a phase shift equal to $\pi/4$ with respect to $\boldsymbol{E}$ and $\boldsymbol{J}$. For $\omega \to +\infty$, the damping tends to infinity and $\boldsymbol{J}=\boldsymbol{E}=\boldsymbol{H}=\mathbf{0}$ for all $z > 0$, so that the conductor becomes impermeable to the electromagnetic field.

From this simple example we infer that the displacement current

$$\frac{\partial \boldsymbol{D}}{\partial t} = \frac{\epsilon}{\gamma}\frac{\partial J(z,t)}{\partial t}\boldsymbol{c}_1 = i\frac{\epsilon\omega}{\gamma} J\boldsymbol{c}_1$$

is negligible with respect to the conduction current $J\boldsymbol{c}_1$ in the conductor provided

$$\omega \ll \frac{\gamma}{\epsilon} \tag{1.80}$$

Besides, some restriction must be added in general on the conductor size. The fact that $\boldsymbol{J}$ is solenoidal implies as we know that any current tube has constant intensity and the current $I = I(t)$ given by the integral (1.6) is constant along the tube for any fixed time t. When applied for example to linear current propagation in a thin wire, this means that the current should propagate instantaneously and should have the same value $I = I(t)$ at all points of the wire. Clearly this cannot be expected to be true if the wire is too long and $\boldsymbol{J}$ varies rapidly with time. In terms of frequency this additional restriction is usually stated as [35]

$$l \ll \frac{c_o}{\omega} \tag{1.81}$$

where l is the conductor length and $c_o = (\epsilon_o\mu_o)^{-1/2}$ is the speed of light in vacuo. Eq. (1.80) is equivalent to $\delta \ll c_o/\omega$, where δ is defined by (1.79), so that (1.81) remains the sole restriction for conductors of length $l > \delta$. In any case we see that the quasi-stationary approximation applies to all ordinary cases of local current conduction in wires [14].

From the mathematical point of view the quasi-stationary Maxwell equations (1.73) are tricky and intriguing, especially from the point of view of the well-posedness of boundary value problems [11].

1.8 Polarization and magnetization.

The electric field $\boldsymbol{E}$ in vacuo equals the ratio $\boldsymbol{D}/\epsilon_o$, whereas in a dielectric or a (bad) conductor $\boldsymbol{D}-\epsilon_o\boldsymbol{E}\neq \boldsymbol{0}$. The deviation is known as the polarization vector $\boldsymbol{P}$:

$$\boldsymbol{D} = \epsilon_o\boldsymbol{E} + \boldsymbol{P} \tag{1.82}$$

If the linear constitutive relation (C2) holds, the polarization vector $\boldsymbol{P}$ in a medium of permittivity ϵ can be defined as

$$\boldsymbol{P} := (\epsilon - \epsilon_o)\boldsymbol{E}$$

The physical interpretation of $\boldsymbol{P}$ goes far beyond the limits of Maxwell's macroscopic model and is related to the formation of electric dipoles induced by the external electric field in the medium [29,35,43]. Since $\epsilon \geq \epsilon_o$, $\boldsymbol{P}$ is directed like $\boldsymbol{E}$ and the polarization opposes the inducing field, in the sense that if $|\boldsymbol{D}|$ is given (i.e. if the charges are kept constant) the intensity of the electric field $|\boldsymbol{E}|$ in the dielectric decreases due to the polarization. Moreover, by definition the polarization $\boldsymbol{P}$ vanishes if the electric field $\boldsymbol{E}$ in the medium vanishes, that is, the polarization is not permanent. This conclusion holds even in the case of polar molecules (like HCl) which possess permanent dipole moments, since in the absence of an external electric field the single dipole moments are oriented at random and hence cancel out. In a perfect conductor we know that $\boldsymbol{D}=\boldsymbol{E}=\boldsymbol{0}$ (see Proposition 1.3.2 above).

[14] For common metals at ordinary temperatures $\gamma \cong 10^7$mho/m, $\mu \cong 10^{-6}$henry/m, $\epsilon \cong 10^{-11}$farad/m, $c_o \cong 3\times 10^8$ m/sec, and so $\omega \ll \gamma/\epsilon \simeq 10^4\,\text{sec}^{-1}$ and l$\cong$ 30 km $\ll c_o/\omega$

Similarly the magnetization vector, defined by

$$\boldsymbol{M} := \frac{1}{\mu_o}\boldsymbol{B} - \boldsymbol{H} \tag{1.83}$$

vanishes in empty space, where $\boldsymbol{B} = \mu_o \boldsymbol{H}$. However, the magnetization can be permanent in certain bodies. Again, the physical explanation of the vector $\boldsymbol{M}$, related to the formation of magnetic dipoles induced in the matter by the external magnetic field $\boldsymbol{B}$, is far outside the limits of Maxwell's model. In non-ferromagnetic bodies, for which the macroscopic constitutive equation (C3) holds, we have

$$\boldsymbol{M} = \frac{\mu - \mu_o}{\mu_o}\boldsymbol{H}$$

and $\boldsymbol{M}$ vanishes if so does $\boldsymbol{H}$. When $\mu > \mu_o$ (paramagnetic bodies) $\boldsymbol{M}$ is directed like $\boldsymbol{H}$ and tends to aid the inducing field $\boldsymbol{B}$, in the sense that if $\boldsymbol{H}$ (e.g. the current in a coil) is kept constant, the field of $\boldsymbol{B}$ is augmented in the body due to magnetization. On the other hand, when $\mu < \mu_o$ (diamagnetic bodies), $\boldsymbol{M}$ is directed like $-\boldsymbol{H}$ and tends to oppose the field of $\boldsymbol{B}$. In both cases, though, the magnetization intensity $|\boldsymbol{M}|$ is very small, so that, as already remarked, in many applications non-ferromagnetic materials can be treated as non-magnetic, i.e. we may take $\mu = \mu_o$.

In contrast, the magnetization $\boldsymbol{M}$ is relevant for ferromagnetic bodies, whose constitutive relation has the form of a nonlinear many-valued function

$$\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{H}) \tag{C3'}$$

with $\boldsymbol{B}$ parallel to $\boldsymbol{H}$, but not necessarily with the same orientation. In particular, if B and H denote the components of $\boldsymbol{B}$ and $\boldsymbol{H}$, respectively, along their common axis, for "magnetically hard" bodies the graphic of $B = B(H)$ has the form of a hysteresis loop (Fig. 1.8). The graphic shows that B and H can have opposite sign and H can vanish while B is different from zero (points 2 and 5 of the figure). Conversely, B can vanish while H is different from zero (points 3 and 6 of the figure). The graphic of the magnetization curve $M = M(H) \equiv B(H)/\mu_o - H$ has a similar shape, and $M(H) \equiv B(H)/\mu_o \neq 0$ for $H = H_2, H = H_5$. The magnetization intensity not only is not a one-valued function of H, but also depends on the previous history of the sample under consideration. If one subjects an initially unmagnetized sample of magnetically hard steel to an increasing magnetic field H, e.g. by using a toroidal coil filled with the material and

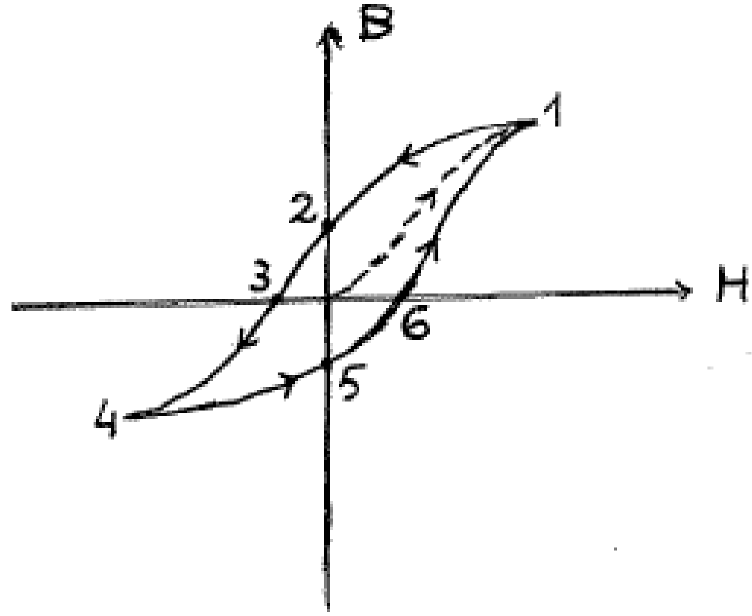

Figure 1.8: Hysteresis loop

increasing the coil current, the initial magnetization curve (shown dotted in the curve) shows an initial sharp rise and subsequently flattens out as a saturation value $M(H_1)$ is reached. If now the applied field is reversed and reduced, the magnetization follows the solid curve $M(H_1) - M(H_4)$. The value $M(H_2)$ for $H = 0$, called residual magnetization or "remanence", and the reversing field H_3 required to reduce M to zero, called coercive force, can be used as measures of the magnetic hardness of the material. If one now carries the magnetization back to the point $M(H_1)$ by increasing H, the lower curve is followed, and the hysteresis loop is completed. This cyclic operation is just what occurs in a-c transformers. The area of the hysteresis loop

$$\oint \boldsymbol{H} \cdot d\boldsymbol{B}$$

is the work per unit volume of the material required to perform a loop, and so there is an energy loss, due to internal friction effects, whenever such a loop is traversed.

In the interior of a permanent magnet the direction of the magnetic field H is generally opposite to that of the magnetization and induction, and the magnetic field intensity can be very small, depending on the geometrical shape of the magnet. Such a magnetization state corresponds to the portion of the magnetization curve lying between $M(H_2)$ and $M(H_3)$. The existence of permanent magnets as opposed to the non-existence of permanent polarization is perhaps the most striking difference between Magnetostatics and Electrostatics.

For magnetically soft bodies, like soft iron, the area enclosed by the loop tends to zero, the loop tends to a single curve, and the nonlinear functions $B = B(H)$, $M = M(H)$ become one-valued (and monotone). As B=B(H) is nonlinear, the permeability μ becomes a function of |H| with a very high average value [15].

The phenomenon of hysteresis is very interesting also from the point of view of mathematical modeling [16].

Appendix:
The MKSA System of Units

In the MKSA system of units, used in these notes [17], the force is expressed in newton, work in joule=newton×meter, power in watt=joule/second, as is well known from mechanics. In electromagnetism a further independent unit is introduced, the ampere (or, equivalently, the coulomb), so that one has four fundamental units:m (meter) for length, kg (kilogram) for mass, sec (second) for time, coulomb=ampere×sec for electric charge. All other electromagnetic units are derived from these four [18].

The charge q is measured in coulomb. From the definitions of electrostatic force $\boldsymbol{F} = q\boldsymbol{E}$ and electric potential $\boldsymbol{E} = -grad\, u$ it follows the electric field $\boldsymbol{E}$ is expressed in newton/ coulomb=volt/m , whereas the potential u as well as the e.m.f. V are measured in volt= joule/coulomb.

Gauss' law implies that $\boldsymbol{D}$ is measured in coulomb/m^2, and the unit of

[15] of the order of 10^4 henry/m

[16] see A.Visintin, Differential Models of Hysteresis, Springer 1994

[17] to be precise, the system adopted here is the so-called rationalized MKSA system due to Giorgi

[18] in principle one could also choose three or five fundamental units. See [43] for an interesting discussion on this issue.

capacity, called farad coincides, by the definition of $\mathbb{C}$, with one coulomb/volt. The permittivity ϵ is then expressed in coulomb×m/(volt×m^2)=farad/m. Eq. (1.22) implies that the unit of magnetic flux Φ, called weber, coincides with one volt×sec, whereas by eq. (1.28) the current I is measures in ampere=coulomb/sec. The magnetic induction $\boldsymbol{B}$ and the current density $\boldsymbol{J}$ are then expressed in weber/m^2 and in ampere/m^2, respectively. Eq. (3.33) implies that the unit of inductance L, called henry, is one weber/ampere, and from Ohm's law it follows that the unit of resistance $\mathcal{R}$ is one ohm=volt/ampere. Eq. (1.35) shows that the conductivity γ is measured in mho/m, where mho= 1/ohm. For common metals $\gamma \simeq 10^7$mhos. Note that one volt×ampere is one watt, the unit of power.

The magnetic intensity field $\boldsymbol{H}$ can be expressed, by force of eq. (3.79), in ampere/m, and so the unit of magnetic permeability μ is one weber $\times m/$ (ampere $\times m^2$) = henry/m. The Poynting vector $\boldsymbol{S}$ is measured in volt×ampere/m^2 =watt/m^2, W_e and W_m are expressed in joule/m^3, and W_J in joule×m^3/sec.

The determination of the remaining (derived) units is left to the reader as an exercise.

Exercises

Exercise 1. (i) For a surface S represented locally in cartesian form $x_3 = f(x_1, x_2)$ the element of area is given by the formula ([19], p. 100)

$$dS = \sqrt{1 + \left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2}\, dx_1 dx_2$$

which extends to surfaces the well-known expression for the arc-length on a curve Γ in the (x_1, x_2)-plane

$$ds = \sqrt{1 + \left(\frac{df}{dx_1}\right)^2}\, dx_1$$

represented locally in cartesian form $x_2 = f(x_1)$.

(ii) Using the Gauss Lemma (LG) yields

$$\int_{\partial\Omega} \boldsymbol{n}\, dS = \int_{\partial\Omega} grad(1)\, dV = \mathbf{0}$$

for every normal domain Ω in $\mathbb{R}^3$.

(iii) Choosing $u = v_k$ in (LG) and summing over all k yields

$$\int_{\partial\Omega} \sum_{k=1}^{3} v_k n_k \, dS \equiv \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} \; dS = \int_{\Omega} \sum_{k=1}^{3} \frac{\partial v_k}{\partial x_k} \mathrm{dx} \equiv \int_{\Omega} div \, \boldsymbol{v} \, \mathbf{dx}$$

Conversely, (LG) follows from (DT) by choosing $v_j = u\delta_{jk}$, where δ_{jk} is the Kronecker delta.

Exercise 2. We have

$$\boldsymbol{n} \cdot curl(x_k \boldsymbol{v}) \equiv x_k \boldsymbol{n} \cdot curl(\boldsymbol{v}) - \boldsymbol{n} \cdot grad \, x_k \wedge \boldsymbol{v} \equiv x_k \boldsymbol{n} \cdot curl \boldsymbol{v} - \boldsymbol{n} \wedge \boldsymbol{v} \cdot \boldsymbol{c}_k$$

$(k = 1, 2, 3)$. The integral of this quantity over $\partial\Omega$ is zero because of Stokes' theorem for a closed surface and since $\boldsymbol{x} = \sum_k x_k \boldsymbol{c}_k$ we obtain the identity

$$\int_{\partial\Omega} \boldsymbol{n} \wedge \boldsymbol{v} \, dS = \int_{\partial\Omega} \boldsymbol{x} \, \boldsymbol{n} \cdot curl \boldsymbol{v} \, dS \qquad (k = 1, 2, 3)$$

This identity is invariant with respect to a shift of the origin of the coordinate system.

Exercise 3. Let Ω be the two-dimensional annular domain $0 < R < |\boldsymbol{x}| < R'$. Show that

$$u = \varphi \equiv \arctan(x_2/x_1)$$

does not satisfy the Gauss Lemma in Ω.

Exercise 4. Let $\boldsymbol{v}$ be regular in Ω. Then:

(i) If $div \, \boldsymbol{v} = 0$ in Ω, the lines of flow of $\boldsymbol{v}$ can be closed or begin and end on $\partial\Omega$ or at infinity.

Hint: $div \, \boldsymbol{v} = 0$ implies that volumes are conserved by the flow map, and

$$\int_S \boldsymbol{v} \cdot \boldsymbol{n} \; dS = \text{constant}$$

for any section S of a current tube. Hence the tubes of flow cannot start or stop at any point in the domain. They may be closed, or start and stop at the boundary, or none of the two. [19]

[19] The popular assertion that streamlines of a solenoidal vector field either are closed or begin and end at the boundary has never been proven.

(*ii*) If *curl* $\boldsymbol{v} = \mathbf{0}$ and Ω is c.s.c., the lines of flow cannot be closed in Ω.

Hint: Suppose Γ is a closed line of flow. By assumption there exists a surface S contained in Ω with $\Gamma = \partial S$, and applying Stokes' theorem (ST1) we obtain

$$\oint_\Gamma \boldsymbol{v} \cdot \boldsymbol{t}\ ds = \int_S curl\, \boldsymbol{v} \cdot \boldsymbol{n}\, dS = 0$$

which contradicts the assumption $|\boldsymbol{v}| \neq 0$ in Ω.

Note that, as in the case of solenoidal fields, the current lines might follow a complicated and possibly chaotical geometrical pattern, wandering endlessly, filling up a region of Ω densely, winding or unwinding with an infinite number of turns without ever closing up.

Exercise 5. Let $\boldsymbol{F}$ be the mechanical force acting on a positive point-like test charge q. As $q \to 0$ we have $\boldsymbol{F} \to \mathbf{0}$. The electric field is defined conceptually as the limit

$$\boldsymbol{E} := \lim_{q \to 0} \frac{\boldsymbol{F}}{q}$$

provided such a limit exists and is finite. In this way the test charge does not perturb the field. If $\boldsymbol{F}$ is proportional to q, as in the Coulomb case, then

$$\boldsymbol{E} = \frac{\boldsymbol{F}}{q}$$

and we obtain eq. (1.1).

Exercise 6. (*i*) The *curl* of a vector field $\boldsymbol{E} = (E_1(\boldsymbol{x}), E_2(\mathbf{x}), E_3(\mathbf{x}))$ is defined in cartesian coordinates (x_1, x_2, x_3) by the expansion according to the elements of the first row of the symbolic determinant

$$curl\, \mathbf{E} = \det \begin{pmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \boldsymbol{c}_3 \\ \partial/\partial x_1 & \partial/\partial x_2 & \partial/\partial x_3 \\ E_1 & E_2 & E_3 \end{pmatrix}$$

and $\boldsymbol{E}$ is irrotational if and only if

$$\frac{\partial E_1}{\partial x_2} = \frac{\partial E_2}{\partial x_1} \quad , \quad \frac{\partial E_3}{\partial x_2} = \frac{\partial E_2}{\partial x_3} \quad , \quad \frac{\partial E_1}{\partial x_3} = \frac{\partial E_3}{\partial x_1}$$

Then

$$\boldsymbol{E} \cdot d\boldsymbol{x} \equiv \sum_{j=1}^{3} E_j dx_j = -du \equiv -\sum_{j=1}^{3} \frac{\partial u}{\partial x_j} dx_j$$

where the potential u is given by

$$u(\boldsymbol{x}) = -\int_{\Gamma} \boldsymbol{E} \cdot d\boldsymbol{x} \; + \mu_0 \;, \quad \mu_0 = u(\boldsymbol{x}_o)$$

Choosing for $\Gamma = \Gamma(\boldsymbol{x}_o,\boldsymbol{x})$ a piecewise straight line we obtain

$$u(\boldsymbol{x}) = -\int_a^{x_1} E_1(y, b, \mathrm{M})\, dy - \int_b^{x_2} E_2(x_1, y, \mathrm{M})\, dy - \int_{\mathrm{M}}^{x_3} E_3(x_1, x_2, y)\, dy + cost.$$

where (a, b, M) is an arbitrary point interior to the domain of definition of $\boldsymbol{E}$.

(*ii*) If the domain is c.s.c. the voltage drop $V = u(\boldsymbol{x}_o) - u(\boldsymbol{x})$ is indipendent of $\Gamma(\boldsymbol{x}_o,\boldsymbol{x})$.

(*iii*) The Coulomb field generated by a source charge (electric pole) $Q = M\epsilon_o$ is irrotational and solenoidal for $r \neq 0$. Hence the Coulomb potential

$$\text{(E1)} \qquad u = \frac{Q}{4\pi\epsilon_o r} \;\equiv\; \frac{M}{4\pi r}$$

is harmonic for $r \neq 0$, i.e. satisfes the Laplace equation

$$\text{(L)} \qquad div\, grad\, \frac{1}{r} \equiv \Delta_3 \frac{1}{r} \equiv \sum_{i=1}^{3} \frac{\partial^2}{\partial x_j^2} \frac{1}{r} \equiv \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) \frac{1}{r} \;=\; 0$$

for $r \neq 0$. For $r \to 0$ the Coulomb potential is singular

$$u(\boldsymbol{x}) = O(1/r) \qquad \text{as } r = |\boldsymbol{x} - \boldsymbol{x}_o| \to 0$$

but $u(\boldsymbol{x})$ is locally integrable in $\mathbb{R}^3$.

In the case of the Coulomb field all lines of force of $\boldsymbol{E}$ are rectilinear and they begin on $\partial\Omega$ and end at infinity (here $\Omega = \mathbb{R}^3 \backslash O$).

Exercise 7. (*i*) Consider two equal and opposite poles of intensity M, a positive one at the point $\boldsymbol{x}_o + \boldsymbol{h}$ and a negative one at $\boldsymbol{x}_o$. A dipole (electric or magnetic) is defined as the conceptual limit

$$\frac{M}{4\pi r'} - \frac{M}{4\pi r} \qquad (r = |\boldsymbol{x}_o - \boldsymbol{x}| \;\boldsymbol{,}\; r\text{'} = |x_o + \boldsymbol{h} - \boldsymbol{x}|)$$

as $|\boldsymbol{h}| \to 0$ and $M \to +\infty$ in such a way that $\boldsymbol{m}= M\boldsymbol{h}$ remains fixed. Let $\widehat{\boldsymbol{h}} = \boldsymbol{h}/h, h = |\boldsymbol{h}|, m = |\boldsymbol{m}|$. Then

$$\boldsymbol{m} = m\widehat{\boldsymbol{h}} \equiv M\boldsymbol{h}$$

is called the moment of the dipole and $\widehat{\boldsymbol{h}}$ its axis. Since

$$\frac{M}{4\pi}\lim_{h\to 0}\left(\frac{1}{r'}-\frac{1}{r}\right) = \frac{m}{4\pi}\ \widehat{\boldsymbol{h}}\cdot grad_{x_o}\frac{1}{r} = \frac{1}{4\pi}\boldsymbol{m}\cdot grad_{x_o}\frac{1}{r} = -\frac{\boldsymbol{m}\cdot(\boldsymbol{x}_o-\boldsymbol{x})}{4\pi r^3}$$

where $r = |\boldsymbol{x}_o - \boldsymbol{x}|$, the dipole potential at the point $\boldsymbol{x}$ is given by

$$v(\boldsymbol{x}) = m\frac{\partial u}{\partial \widehat{\boldsymbol{h}}} \equiv \frac{\boldsymbol{m}\cdot(\boldsymbol{x}-\boldsymbol{x}_o)}{4\pi r^3}$$

where u is the source potential (E1).

(ii) Verify that $v(\boldsymbol{x})$ is harmonic for $r \neq 0$ and that the field due to a (say, magnetic) dipole $\boldsymbol{H} = -\mathrm{grad}\, v$ has the cartesian components

$$H_j \equiv -\frac{\partial v}{\partial x_j} = -\frac{1}{4\pi r^5}\sum_{i=1}^{3} r_i(m_j r_i - 3m_i r_j) \qquad (j = 1,2,3)$$

Exercise 8. (compass needle). Let θ denote the angle between $\mathbf{m}$ and the Earth magnetic field $\boldsymbol{B}_o$. Then $\boldsymbol{T}=\boldsymbol{m}\wedge\boldsymbol{B} = -mB_o sin\theta\,\mathbf{k}$, where $\mathbf{k}$ is the unit vector orthogonal to the plane $(\boldsymbol{m},\boldsymbol{B}_o)$. The equation of motion of the compass needle is

$$\mathcal{I}\frac{d^2\theta}{dt^2} = -mB_o sin\theta$$

where $\mathcal{I}$ is the moment of inertia. This is the equation for the oscillations of a pendulum, having the position $\theta = 0$ as a stable equilibrium, and in correspondence $\boldsymbol{m}$ and $\boldsymbol{B}_o$ have the same direction. Since in practice a small friction term $-\varepsilon d\theta/dt$ $(\varepsilon > 0)$ must always be added on the right-hand side of the above equation, this equilibrium becomes asymptotically stable and all solutions $\theta(t)$ tend to zero as t tends to infinity.

Exercise 9. Take a coordinate system with origin at the center of the circular loop and the axes chosen in such a way that $\mathbf{b}=\mathbf{c}_3$ and $\boldsymbol{B} = B_1\mathbf{c}_1 + B_3\mathbf{c}_3$. From eq. (1.14) we obtain

$$\oint \boldsymbol{x}\wedge \boldsymbol{dF} = I\oint \boldsymbol{x}\wedge(\boldsymbol{t}\wedge\boldsymbol{B})ds \equiv I\oint \boldsymbol{x}\cdot\boldsymbol{B}\ \mathbf{t}ds - I\boldsymbol{B}\oint \boldsymbol{x}\cdot\frac{d\boldsymbol{x}}{ds}ds = I\oint \boldsymbol{x}\cdot\boldsymbol{B}\ \mathbf{t}ds$$

since

$$\oint \boldsymbol{x}\cdot\frac{d\boldsymbol{x}}{ds}\,ds=\frac{1}{2}\oint\frac{d|\boldsymbol{x}|^2}{ds}\,ds=0$$

In polar coordinates ϱ,φ we have

$$\boldsymbol{x}\cdot\boldsymbol{B}=x_1B_1=B_1R\,cos\varphi,\;ds=R\,d\varphi,\;\boldsymbol{t}=cos\varphi\mathbf{c}_2-sin\varphi\mathbf{c}_1\,,\;\mathbf{b}\wedge\boldsymbol{B}=B_1\mathbf{c}_2$$

Hence

$$\boldsymbol{T}=I\oint\boldsymbol{x}\cdot\boldsymbol{B}t\,ds=IB_1R\int_0^{2\pi}cos\varphi(cos\varphi\mathbf{c}_2-sin\varphi\mathbf{c}_1)R\,d\varphi=I\pi R^2B_1\mathrm{c}_2=I\mathcal{A}\mathbf{b}\wedge\mathbf{B}$$

Exercise 10.(The Biot-Savart law (1.19)). Suppose that the wire coincides with the $z-$axis, so $s=x_3$, $\boldsymbol{t}$=$\mathbf{c}_3$, $\boldsymbol{t}\wedge\boldsymbol{r}$=$\mathbf{c}_3\wedge\boldsymbol{r}$=$\mathbf{c}_3\wedge\varrho\mathbf{u}$= $\varrho\boldsymbol{\tau}$, where $\boldsymbol{u}$ is the radial unit vector in the plane orthogonal to the wire. Letting ψ =arcsin(ϱ/r) =arctan$\,(s/\varrho)$ we obtain

$$\frac{ds}{r^3}=\frac{\varrho\,d\psi}{r^3cos^2\psi}=\frac{cos\psi\,d\psi}{\varrho^2}\;,\quad\frac{I}{4\pi}\int_{-\infty}^{+\infty}\frac{\mathbf{t}\wedge\mathbf{r}}{r^3}\,ds=\frac{I\boldsymbol{\tau}}{4\pi\varrho}\cdot\int_{-\pi/2}^{\pi/2}cos\psi\,d\psi=\frac{I\boldsymbol{\tau}}{2\pi\varrho}$$

where $\boldsymbol{\tau}:=-sin\varphi\mathbf{c}_1+cos\varphi\mathbf{c}_2=\varrho grad\,\varphi$. Therefore $\boldsymbol{H}=-\mathrm{grad}\,v$ is a one-valued vector field, irrotational and solenoidal for $\varrho>0$. Since the two-dimensional Laplace operator Δ_2 in polar coordinates has the expression

$$\Delta_2u=u_{x_1x_1}+u_{x_2x_2}=u_{\varrho\varrho}+\frac{1}{\varrho}u_{\varphi\varphi}$$

φ is a many-valued harmonic function for $\varrho>0$, and a branch surface for φ is any arbitrary plane φ =constant. Since any two branches of φ differ by a multiple of 2π, the potential $v(\boldsymbol{x})=-\frac{I}{2\pi}\varphi$ is a many-valued harmonic function for $\varrho>0$ with period $-I$ and with the same branch surfaces φ =constant.

Exercise 11. Consider a homogeneous dielectric of permittivity ϵ subjected to a constant electric field $\boldsymbol{E}_o$. By carving a long thin cavity in the dielectric, the field in the center of the cavity is the same as $\boldsymbol{E}_o$ if the cavity is parallel to $\boldsymbol{E}$, is given by $\epsilon\boldsymbol{E}/\epsilon_o$ if it is orthogonal.

Hint: use the matching relations (R1) and (R5).

Exercise 12 (Proof of (1.42). Introducing a system of local cartesian coordinates (x_1, x_2, x_3) in the neighborhood of a generic point $\boldsymbol{y}$ of $\mathbb{S}$, with x_3-axis parallel to the normal $\boldsymbol{n}(\boldsymbol{y})$, the relation $[\boldsymbol{n}\wedge grad\, u] = \boldsymbol{0}$ becomes

$$[\frac{\partial u}{\partial x_j}] := (\frac{\partial u}{\partial x_j})_+ - (\frac{\partial u}{\partial x_j})_- = 0 \qquad (j = 1, 2)$$

But by the definition of the jump

$$(\frac{\partial u}{\partial x_j})_+ - (\frac{\partial u}{\partial x_j})_- = \frac{\partial u_+}{\partial x_j} - \frac{\partial u_-}{\partial x_j} := \frac{\partial [u]}{\partial x_j} \qquad (j = 1,\ 2)$$

Hence $[\boldsymbol{n}\wedge grad\, u] = \boldsymbol{n}\wedge grad\, [u] = \boldsymbol{0}$, and (1.42) follows.

Exercise 13 (Joule's law). For a wire of length l , section area A, conductivity γ , resistance $\mathcal{R}$ given by eq. (1.35) electromotive force V and carrying the current I, we have

$$\frac{1}{\gamma} = \frac{\mathcal{R}A}{l} \quad , |\boldsymbol{J}| = \frac{I}{A}$$

and therefore

$$P_J = \gamma^{-1}|\boldsymbol{J}|^2 = \frac{I^2}{\gamma A^2} = \frac{\mathcal{R}I^2}{lA}$$

Since lA is the volume of the wire, the power dissipated into heat is

$$P = \mathrm{V}I \equiv \mathcal{R}I^2$$

Exercise 14. Consider a point charge Q concentrated at an interior point $\boldsymbol{x}_o$ of a bounded domain Ω filled by a homogeneous dielectric. Then $\boldsymbol{E}(\boldsymbol{x})$ is given by the Coulomb law

$$\boldsymbol{E}(\boldsymbol{x}) = \frac{Q}{4\pi\epsilon r^2} \frac{(\boldsymbol{x} - \boldsymbol{x}_o)}{|\boldsymbol{x} - \boldsymbol{x}_o|}$$

and the integral expressing the local electric energy (1.50)

$$\mathcal{E}_\Omega(t) := \int_\Omega \frac{1}{2}\epsilon|\boldsymbol{E}(\boldsymbol{x}, t)|^2\, \boldsymbol{dx} = \frac{Q^2}{32\epsilon\pi^2} \int_\Omega \frac{1}{|\boldsymbol{x} - \boldsymbol{x}_o|^4} \boldsymbol{dx}$$

is divergent (see Exercise 19).

Similarly the magnetic field due to a magnetic dipole of moment $\boldsymbol{m}$ concentrated at $\boldsymbol{x}_o$ has a singularity of order $O(|\boldsymbol{x}-\boldsymbol{x}_o|^{-3})$ as $\boldsymbol{x}\to\boldsymbol{x}_o$. Therefore the integral for the local magnetic energy

$$\mathcal{E}_\Omega(t) := \int_\Omega \frac{1}{2}\mu|\boldsymbol{H}(\boldsymbol{x},t)|^2\, \boldsymbol{dx} = \int_\Omega O(|\boldsymbol{x}-\boldsymbol{x}_o|^{-6})\, \boldsymbol{dx}$$

diverges.

Exercise 15. We have using Stokes' theorem,

$$\int_{\partial\Omega} u\boldsymbol{J}\cdot\boldsymbol{n}\, dS = \int_{\partial\Omega} u\boldsymbol{n}\cdot curl\boldsymbol{H}\, dS \equiv \int_{\partial\Omega} \boldsymbol{n}\cdot curl(u\boldsymbol{H})\, dS - \int_{\partial\Omega} \boldsymbol{n}\cdot grad\, u\wedge\boldsymbol{H}\, dS$$

$$\equiv \int_{\partial\Omega} \boldsymbol{E}\wedge\boldsymbol{H}\cdot\boldsymbol{n}\, dS$$

Exercise 16. Let Ω be the s.m.c. domain $R < |\boldsymbol{x}| < R'$ bounded by two concentric spheres. For arbitrary real constants $a \neq 0$ and b, the non-zero vector field

$$\boldsymbol{v} = -grad\,(\frac{a}{|\boldsymbol{x}|} + b)$$

is irrotational and solenoidal for $|\boldsymbol{x}| > 0$ and satisfies the boundary condition $\boldsymbol{v}\wedge\boldsymbol{n}\equiv \boldsymbol{0}$ for $|\boldsymbol{x}| = R$ and $|\boldsymbol{x}| = R'$. The scalar potential $u = a/|\boldsymbol{x}|+b$ is clearly one-valued in Ω.

Exercise 17. $\boldsymbol{E}= -grad\,u, \boldsymbol{H}= -grad\,v$ with potentials u, v that depend solely on $r = |\boldsymbol{x}|$ for symmetry reasons and are one-valued as C and D are contourwise simply connected. It follows that $\boldsymbol{E}, \boldsymbol{H}$ are radial and $\boldsymbol{E}\wedge\boldsymbol{n}=\boldsymbol{H}\wedge\boldsymbol{n}= \boldsymbol{0}$ for $r = R$. Lemma 1.6.1(i) shows that $\boldsymbol{E}=\boldsymbol{H}= \boldsymbol{0}$ for $r < R$, so that $\boldsymbol{B}= \boldsymbol{0}$ and $\boldsymbol{H}\cdot\boldsymbol{n}= 0$ for $r = R$. Lemma 1.6.2(ii) implies then that $\boldsymbol{H}= \boldsymbol{0}$ everywhere. The electric potential u is a radial harmonic function for $r > R$, hence $u = b+\mathrm{M}/r$. Moreover by (1.66)

$$-\frac{du}{dr} = \frac{\sigma_o}{\epsilon} \qquad \text{for } r = R$$

and so $\mathrm{M}= \sigma_o R^2/\epsilon$. Hence

$$u = \frac{R^2\sigma_o}{\epsilon|\boldsymbol{x}|} \qquad \text{for } r > R$$

and $\boldsymbol{E}= -grad\, u$ is the unique solution satisfying (1.54). The potential u is the so-called "capacitary potential" of the sphere $r = R$.

Exercise 18. Since $(1+i)^2 = 2i$, it is immediate to see that

$$J(z,t) = Ae^{i\omega t} exp\big[-(1+i)\frac{z}{\delta})\big] \ , \qquad \delta := \sqrt{\frac{2}{\gamma\mu\omega}}$$

satisfies the boundary condition at $z = 0$ and is a bounded solution of the equation

$$\frac{\partial J}{\partial t} = \frac{2}{\omega\delta^2}\frac{\partial^2 J}{\partial z^2} \equiv \frac{1}{\mu\gamma}\frac{\partial^2 J}{\partial z^2}$$

for $z > 0$. Then $\boldsymbol{E}=\boldsymbol{J}/\gamma$ and by the first equation (1.73)

$$\mu_o\frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E} \ \Rightarrow\ \boldsymbol{H} = \frac{1}{i\omega\mu_o}\frac{\partial E}{\partial z}\,\mathbf{c}_2 \ = \frac{(1+i)\ J(z,t)}{i\omega\mu_o\gamma\delta}\,\mathbf{c}_2$$

where $(1+i)/i = 1 - i = \sqrt{2}e^{-i\pi/4}$ and $\omega\mu_o\gamma\delta = \sqrt{2\omega\mu_o\gamma}$.

The displacement current is negligible with respect to the conduction current in the conductor if $\epsilon|\partial\boldsymbol{E}/\partial t| \equiv \epsilon\gamma^{-1}|\partial\boldsymbol{J}/\partial t| \equiv \epsilon\gamma^{-1}\omega|\boldsymbol{J}| \ll |\boldsymbol{J}|$, that is if $\gamma^{-1}\omega \ll 1/\epsilon$.

Exercise 19. If K is a bounded domain of $\mathbb{R}^n$ $(n \geq 1)$, the integral

$$\int_K \frac{1}{|\boldsymbol{x}-\boldsymbol{x}_o|^\alpha}\mathrm{d}\boldsymbol{x} \qquad\qquad (\boldsymbol{x}_o \in K)$$

exists finite if $\alpha < n$, diverges for $\alpha \geq n$.

Hint: for a ball of center $\boldsymbol{x}_o$ one has $|\boldsymbol{x}-\boldsymbol{x}_o|^{-\alpha}d\mathbf{x} = r^{n-1-\alpha}drd\Omega$.

Chapter 2

Electrostatics and Magnetostatics

We study in this chapter the electrostatic field generated by a charge or a system of charges embedded in a single homogeneous isotropic dielectric medium D and their effects on one or several homogeneous isotropic conductors C. The results of Chapter 1 and in particular Corollary 1.6.10 show that the presence of electric charge in some shape is necessary in order to have a non-vanishing electrostatic field. We recall from §1.6.2 of Chapter 1 that the fundamental unknown in Electrostatics is the electric potential $u = u(\boldsymbol{x})$, which satisfies the Poisson equation

$$\Delta_3 u = -\frac{\rho(\boldsymbol{x})}{\epsilon} \tag{2.1}$$

for $\boldsymbol{x} \in D$, where ϵ is the constant permittivity of D, and $\rho(\boldsymbol{x})$ is the volume charge density. It follows that $u(\boldsymbol{x})$ is a harmonic function, i.e. satisfies the Laplace equation

$$\Delta_3 u(\boldsymbol{x}) = 0 \tag{2.2}$$

in any open set Ω where $\rho(\boldsymbol{x}) \equiv 0$, in particular in a conductor. Actually, since the current density $\boldsymbol{J}$ is everywhere zero by assumption, Ohm's law $\boldsymbol{J} = \gamma \boldsymbol{E}$ implies that inside a conductor the electric field $\boldsymbol{E}(\boldsymbol{x}) = -grad\, u(\boldsymbol{x})$ is identically zero and so the potential is constant.

Consider a conductor C surrounded by a dielectric D, and let $\boldsymbol{n}$ denote the normal to ∂C, oriented from C ($-$ side) towards D ($+$ side) [1]. Then

$$u(\boldsymbol{x}) = u_{\_} \equiv \text{constant}$$

for $\boldsymbol{x} \in C$, and by continuity the inner limit of $u(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches ∂C on the conductor side is given by the same constant u_- . We denote by

$$u_+ := \lim_{\substack{\boldsymbol{x}\to\partial C \\ \boldsymbol{x}\in D}} u(\boldsymbol{x})$$

the limit of u from $+$ side, which represents the trace $u|_{\partial C}$ of u over the conductor surface as "seen" from the dielectric D surrounding it. Then

$$u|_{\partial C} \equiv u_+ = u_{\_} + [u]_{\partial C}$$

where the jump $[u]_{\partial C}$ is constant on every connected component of ∂C (see §1.4). If the potential is continuous across ∂C, then $[u]_{\partial C} = 0$ and the trace $u|_{\partial C} = u_{\_}$ is constant over all the boundary ∂C, whether connected or not.

We have seen in Chapter 1 that the surface charge density on a surface $\mathbb{S}$,with $\epsilon_+ = \epsilon_- = \epsilon$, is given by

$$\sigma = [\boldsymbol{D}]_{\mathbb{S}} \cdot \boldsymbol{n} = \epsilon\,(\boldsymbol{E}_+ \cdot \boldsymbol{n} - \boldsymbol{E}_{\_} \cdot \boldsymbol{n}) \tag{2.3}$$

Here, since $\boldsymbol{E}(\boldsymbol{x}) \equiv \boldsymbol{0}$ in the conductor, $\boldsymbol{E}_{\_}\cdot\boldsymbol{n} = 0$ and the surface charge density at the point $\boldsymbol{y}$ on the conductor surface ∂C is given by the limit

$$\sigma(\boldsymbol{y}) = \epsilon \boldsymbol{E}_+(\boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{y}) \equiv -\epsilon \lim_{\boldsymbol{x}\to\boldsymbol{y}} \boldsymbol{n}(\boldsymbol{y}) \cdot grad\, u(\boldsymbol{x}) \qquad (\boldsymbol{y} \in \partial C) \tag{2.4}$$

as $\boldsymbol{x}\to\boldsymbol{y}$, $\boldsymbol{x} \in D$. Moreover, as $\boldsymbol{E}\wedge\boldsymbol{n} = \boldsymbol{0}$ on ∂C, the lines of force of $\boldsymbol{E}$ (and of $\boldsymbol{D} = \epsilon\boldsymbol{E}$) in the dielectric are orthogonal to the conductor surface.

We will assume as a rule (with exceptions) that the potential $u(\boldsymbol{x})$ is regular at infinity, i.e.that it satisfies the asymptotic condition (1.54) of Chapter I

$$u(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-1}) \quad , \quad grad\, u(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-2}) \qquad \text{as } |\boldsymbol{x}| \to \infty \tag{2.5}$$

uniformly with respect to direction [2]. We will also assume that all conductors and dielectrics under consideration are rigid bodies.

[1] we recall that by assumption C, D, ... represent open connected sets (domains) in $\mathbb{R}^3$

[2] there exist constants $M, R > 0$ such that $|\boldsymbol{x}||u(\boldsymbol{x})| < M$ for all $|\boldsymbol{x}| > R$, and similarly for $grad\, u$

Remark 1. Suppose that the dielectric D is bounded and the conductor $C = \mathbb{R}^3 \backslash \overline{D}$ is unbounded. Then if $\boldsymbol{n}$ is the inner normal on ∂D, in accordance with the orientation chosen above, the divergence theorem shows that

$$(2.6) \qquad - \int_D \Delta_3 \, u \; dV \; = \int_{\partial D} \frac{\partial u}{\partial n} dS$$

where $\partial u / \partial n$ is the normal trace of $grad\, u$ in D :

$$(2.7) \qquad \frac{\partial u(\boldsymbol{y})}{\partial n} \;\; := \;\; \lim_{\boldsymbol{x} \to \boldsymbol{y}} \boldsymbol{n}(\boldsymbol{y}) \cdot grad\, u(\mathbf{x})$$

and the limit is defined as in (2.4) ($\boldsymbol{x} \in D$, $\boldsymbol{y} \in \partial C$). In this case since $u = u_-$ is constant in the conductor $C = \mathbb{R}^3 \backslash \overline{D}$,

$$(2.8) \qquad \frac{\partial u}{\partial n} = (\frac{\partial u}{\partial n})_+ \quad , \quad (\frac{\partial u}{\partial n})_- = 0$$

From eqs. (2.1) and (2.4) we then obtain

$$(2.9) \qquad \int_D \rho \, dV \;\; = - \int_{\partial D} \sigma \; dS$$

This equation says that the total volume charge in D is equal and opposite to the total surface charge on ∂D, given by the the left-hand side of eq. (2.9).

In contrast, eqs. (2.6) and (2.9) do not hold in general in the more common situation of a bounded conductor C surrounded by an unbounded dielectric $D = \mathbb{R}^3 \backslash \overline{C}$, if u is regular at infinity (Exercise 1).

From the mathematical point of view, this chapter is about potential theory and some properties of solutions of elliptic partial differential equations of which those of Laplace and Poisson are prototypes [2,18,37].

2.1 Electrostatic field in a dielectric

The simplest problem of Electrostatics is the

Summation problem: find the electrostatic potential $u(\boldsymbol{x})$ due to a given charge distribution, discrete or continuous, in a single unbounded dielectric $D = \mathbb{R}^3$ having constant permittivity ϵ.

The case of a single point charge Q concentrated at the point $\boldsymbol{y} \in \mathbb{R}^3$ is solved by Coulomb's law (1.38) : the potential u and the electric field $\boldsymbol{E} = -grad\, u$ at the point $\boldsymbol{x} \neq \boldsymbol{y}$ are given by

$$u(\boldsymbol{x}) = \frac{Q}{4\pi\epsilon r}\,,\ \boldsymbol{E}(\boldsymbol{x}) = \frac{Q}{4\pi\epsilon r^2}\frac{\boldsymbol{x}-\boldsymbol{y}}{r} \tag{2.10}$$

where $r = |\boldsymbol{x}-\boldsymbol{y}|$. We next consider the case of an electric dipole at the point $\boldsymbol{y}$ with moment $\boldsymbol{m}$. We know (see Exercise 7 of Chapter 1) that the potential and the electric field at the point $\boldsymbol{x} \neq \boldsymbol{y}$ due to such a dipole are given by

$$u(\boldsymbol{x}) = \frac{\boldsymbol{m}\cdot(\boldsymbol{x}-\boldsymbol{y})}{4\pi r^3},\ E_j = -\frac{1}{4\pi r^5}\sum_{i=1}^{3} r_i(m_j r_i - 3m_i r_j)\ (j = 1,2,3) \tag{2.11}$$

The corresponding charge volume density $\rho(\boldsymbol{x})$ is the Dirac distribution $Q\delta_{\boldsymbol{y}}(\boldsymbol{x})$ in the case of the point charge and the derivative of the Dirac distribution $\boldsymbol{m}\cdot grad\, \delta_{\boldsymbol{y}}(\boldsymbol{x})$ in the case of the dipole. In both cases $u(\boldsymbol{x})$ is harmonic for all $\boldsymbol{x} \neq \boldsymbol{y}$ and regular at infinity.

In the case of continuous charge distributions with given smooth volume density $\rho(\boldsymbol{x})$ or surface density $\sigma(\boldsymbol{x})$ one has to solve a boundary value problem for the Poisson or Laplace equation (see Exercises 1, 2). A summation problem which can be solved easily using the Gauss Law (1.4) is that of a uniform surface charge distribution σ over a homogeneous conducting plane. Suppose that the half−space $z < 0$ is conducting and $z > 0$ is a dielectric medium with constant permittivity ϵ. If $\boldsymbol{n} = \mathbf{c}_3$ is the unit normal to the separation surface, the electric field $\boldsymbol{E}$ and the displacement vector $\boldsymbol{D} = \epsilon\boldsymbol{E}$ vanish in the conductor $z < 0$ and are given by the constant vector fields

$$\boldsymbol{E} = \frac{\sigma}{\epsilon}\boldsymbol{n} \quad , \quad \boldsymbol{D} = \sigma\boldsymbol{n} \tag{2.12}$$

in the dielectric $z > 0$ (Exercises 3 and 4). The corresponding electrostatic potential

$$u = -\frac{\sigma}{\epsilon}z + ç \tag{2.13}$$

is constant ($u =$ç) over the separation surface $z = 0$. Precisely, $u_+ =$ç is the boundary value "seen from the dielectric" $z > 0$, $u_- =$ç$-[u]_{\partial C}$ is the constant value in the conductor $z < 0$, where $[u]_{\partial C}$ is the potential jump (if any).

This result can be extended locally to an arbitrary homogeneous conductor with smooth boundary surface S: S is equipotential and eqs. (2.12), (2.13) hold in the neighborhood of every regular point $\mathbf{y}$ of S with z a local normal coordinate [3] and σ, $\mathbf{n}$ depending in general on $\mathbf{y}$.

2.1.1 Volume potentials and single layer potentials.

We next consider a volume distribution of charge with density $\rho(\boldsymbol{y})$, a bounded function of compact support $K \subset \mathbb{R}^3$. By the superposition principle, the potential at the point $\boldsymbol{x}$ is expected to be given by the integral over $\boldsymbol{y} \in K$ of the Coulomb potentials (2.10) with $Q = \rho \boldsymbol{dy}$

$$(2.14) \qquad \mathbb{V}(\boldsymbol{x}) := \frac{1}{4\pi} \int_K \frac{\rho(\boldsymbol{y})}{\epsilon} \frac{\boldsymbol{dy}}{|\boldsymbol{x}-\boldsymbol{y}|} \equiv \frac{1}{4\pi\epsilon} \int_{\mathbb{R}^3} \frac{\rho(\boldsymbol{y})\, \boldsymbol{dy}}{|\boldsymbol{x}-\boldsymbol{y}|}$$

This can in fact be proven:under reasonable assumptions (see Theorem 1) the function $\mathbb{V} = \mathbb{V}_{\rho/\epsilon}(\boldsymbol{x})$, called volume potential with density ρ/ϵ, is a solution of Poisson's equation (2.1) satisfying (2.5). More precisely,

$$(2.15) \qquad \mathbb{V}_{\rho/\epsilon}(\boldsymbol{x}) \sim \frac{1}{4\pi\epsilon|\boldsymbol{x}|} \int_K \rho(\boldsymbol{y})\, \boldsymbol{dy} \quad + \quad O(|\boldsymbol{x}|^{-2})$$

as $|\boldsymbol{x}| \to +\infty$ (Exercise 5 and 25).

Theorem 2.1.1 (*i*) *If $\rho(\boldsymbol{y})$ is bounded with compact support $K \subset \mathbb{R}^3$, $\mathbb{V}$ is a $C^1(\mathbb{R}^3)$ bounded function with bounded gradient and satisfies (2.15).*

(*ii*) *If in addition $\rho \in C^1(K) \Rightarrow \mathbb{V} \in C^1(\mathbb{R}^3) \cap C^2(\mathbb{R}^3 \backslash \partial K)$ is a bounded solution of (1) in $\mathbb{R}^3 \backslash \partial K$ and is regular at infinity.*

(*iii*) *If in addition $\rho \in C^1(\mathbb{R}^3) \Rightarrow \mathbb{V} \in C^2(\mathbb{R}^3)$ satisfies (1) in all $\mathbb{R}^3$.*

(For the proof, see e.g. [2].)

---
[3] z is a local coordinate "adapted" to S

Similarly, the principle of superposition suggests that the electrostatic potential at the point $\boldsymbol{x} \notin \mathbb{S}$ due to a surface distribution of charge with density $\sigma(\boldsymbol{y})$ over a regular bounded surface $\mathbb{S} \subset D$ (open or closed), is given by the surface integral

$$\mathcal{V}(\boldsymbol{x}) := \int_{\mathbb{S}} \frac{\sigma(\boldsymbol{y})\, dS_y}{4\pi\epsilon r} \quad , \qquad r = |\boldsymbol{x} - \boldsymbol{y}| \tag{2.16}$$

denoted also $\mathcal{V}_{\sigma/\epsilon}$ and called single layer potential with density σ/ϵ. If σ is bounded $\mathcal{V}_{\sigma/\epsilon}$ satisfies the asymptotic condition (2.5): more precisely we have

$$\mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x}) \sim \frac{1}{4\pi\epsilon|\boldsymbol{x}|} \int_{\mathbb{S}} \sigma(\boldsymbol{y})\, dS \;+\; O(|\boldsymbol{x}|^{-2}) \qquad \text{as} \;\; |\boldsymbol{x}| \to +\infty \tag{2.17}$$

(Exercise 5).

Theorem 2.1.2 *[2]. If $\sigma \in C^1(\mathbb{S})$ and $\mathbb{S}$ is an orientable bounded C^2 surface* [4]*, then:*

(a) $\mathcal{V} = \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x})$ is bounded and continuous in $\mathbb{R}^3$

(b) $\mathcal{V} \in C^\infty(\mathbb{R}^3 \backslash \mathbb{S})$ is harmonic in $\mathbb{R}^3 \backslash \mathbb{S}$ and satisfies (2.17)

(c) $grad\, \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x})$ is continuous for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \overline{\mathbb{S}}$ and satisfies the jump relations

$$[grad\, \mathcal{V}]_{\mathbb{S}} := (grad\, \mathcal{V}_{\sigma/\epsilon})_+ - (grad\, \mathcal{V}_{\sigma/\epsilon})_- = -\frac{\sigma(\boldsymbol{y})}{\epsilon}\boldsymbol{n}(\boldsymbol{y}) \tag{2.18}$$

at all points $\boldsymbol{y} \in \mathbb{S}$.

In other words, the normal derivative of the single layer potential is discontinuous, whereas the tangential derivatives are continuous, at every point of $\mathbb{S}$:

$$\begin{cases} [\frac{\partial \mathcal{V}}{\partial n}]_{\mathbb{S}} := (\boldsymbol{n} \cdot grad\, \mathcal{V}_{\sigma/\epsilon})_+ - (\boldsymbol{n} \cdot grad\, \mathcal{V}_{\sigma/\epsilon})_- = -\frac{\sigma}{\epsilon} \\ [\mathbf{n} \wedge grad\, \mathcal{V}]_{\mathbb{S}} = \mathbf{0} \end{cases} \tag{2.19}$$

Since by assumption ϵ is constant everywhere, Theorem 2.1.2 implies that the electrostatic field $\boldsymbol{E}(\boldsymbol{x}) = -grad\mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x})$ and the displacement vector $\boldsymbol{D} = \epsilon\boldsymbol{E}$ are irrotational and solenoidal smooth vector fields for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \mathbb{S}$,

[4] these assumptions can be relaxed [2]

and their ± limits exist on $\mathbb{S}$ and satisfy the matching relations at every point of $\mathbb{S}$

$$[\boldsymbol{D}]_{\mathbb{S}} \cdot \boldsymbol{n} = \sigma \quad , \quad [\boldsymbol{E}]_{\mathbb{S}} \wedge \boldsymbol{n} = \boldsymbol{0}$$

in accordance with the results of §1.4.

Note that, if $\mathbb{S}$ is open, $\boldsymbol{E} = -grad\, \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x})$ has a logarithmic singularity at the points of $\partial\mathbb{S}$ where $\sigma \neq 0$ (Exercise 6 and 7).

2.1.2 Double layer potentials.

We have repeatedly pointed out that in the Electrostatics of conductors we need include also potentials that are bounded and piecewise continuous, with jump discontinuities $[u]_{\partial C}$ that are constant over any connected component $\mathbb{S}$ of the conductor boundary ∂C. This can be obtained by means of a surface distribution of electric dipoles (2.11) with moment $\boldsymbol{m} = \nu \boldsymbol{n}\, dS$ and axis $\widehat{\mathbf{h}} = \boldsymbol{n}$, the unit normal to $\mathbb{S}$

$$\mathcal{W}_\nu(\boldsymbol{x}) = \frac{1}{4\pi} \int_{\mathbb{S}} \nu(\boldsymbol{y}) \frac{\partial}{\partial n_y} \left(\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|}\right) dS_{\boldsymbol{y}}$$

This integral is called double layer potential with density $\nu(\boldsymbol{y})$. The case of interest here is $\nu(\boldsymbol{y}) = \nu_o$ =constant on $\mathbb{S} = \partial C$ connected:

$$\mathcal{W}_{\nu_o}(\boldsymbol{x}) = \frac{\nu_o}{4\pi} \int_{\mathbb{S}} \frac{\partial}{\partial n_y} \left(\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|}\right) dS_{\boldsymbol{y}} \tag{2.20}$$

Clearly

$$\mathcal{W}_{\nu_o}(\boldsymbol{x}) \equiv \nu_o \mathcal{W}_1(\boldsymbol{x})$$

where $\mathcal{W}_1(\boldsymbol{x})$ denotes the double layer potential with unit density.

Theorem 2.1.3 *[2]. If $\mathbb{S}$ is a bounded connected C^2 surface, then*

(i) $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is harmonic for $\boldsymbol{x} \in \mathbb{R}^3 \setminus \mathbb{S}$, is continuous for $\boldsymbol{x} \in \mathbb{S}$, and is discontinuous as $\boldsymbol{x}$ crosses $\mathbb{S}$.

(ii) $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is bounded for $\boldsymbol{x} \in \mathbb{R}^3$ and vanishes (uniformly) at infinity:

$$\mathcal{W}_{\nu_o}(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-2}) \qquad as\ |\boldsymbol{x}| \to +\infty$$

(iii) The limits $\mathcal{W}_+(\boldsymbol{y})$, $\mathcal{W}_-(\boldsymbol{y})$ of $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ exist as $\boldsymbol{x}$ approaches a point $\boldsymbol{y}\in\mathbb{S}$ on the two sides of $\mathbb{S}$, and the jump relations hold

$$[\mathcal{W}_{\nu_o}]_{\mathbb{S}} := \mathcal{W}_+(\boldsymbol{y}) - \mathcal{W}_-(\boldsymbol{y}) = \nu_o \qquad \boldsymbol{y}\in\mathbb{S} \tag{2.21}$$

with a potential jump ν_o constant along $\mathbb{S}$.

(iv) As $\boldsymbol{y}$ varies over $\mathbb{S}$ we have $\mathcal{W}_+(\boldsymbol{y})\in C^o(\mathbb{S})$, $\mathcal{W}_-(\boldsymbol{y})\in C^o(\mathbb{S})$ and

$$\frac{1}{2}(\mathcal{W}_+(\boldsymbol{y}) + \mathcal{W}_-(\boldsymbol{y})) = \nu_o \int_{\mathbb{S}} \frac{\partial}{\partial n_y}\Big(\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}\Big)\, dS_y \;\equiv\; \mathcal{W}_{\nu_o}(\boldsymbol{y}) \tag{2.22}$$

(v) If $\mathbb{S}$ is flat we have $\mathcal{W}_{\nu_o}(\boldsymbol{y})\equiv 0\ \ \forall\boldsymbol{y}\in\mathbb{S}$ (Exercise 8).

(vi) The gradient of $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is continuous in $\mathbb{R}^3\backslash\overline{\mathbb{S}}$ and satisfies the continuous matching relation across $\mathbb{S}$

$$[grad\,\mathcal{W}_{\nu_o}]_{\mathbb{S}} \equiv (grad\,\mathcal{W}_{\nu_o})_+ - (grad\,\mathcal{W}_{\nu_o})_- = \boldsymbol{0} \tag{2.23}$$

Hence by defining the gradient of $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ on $\mathbb{S}$ by continuity

$$grad\,\mathcal{W}_{\nu_o}(\boldsymbol{y}) := (grad\,\mathcal{W}_{\nu_o})_+ \qquad \boldsymbol{y}\in\mathbb{S}$$

$grad\,\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is continuous in $\mathbb{R}^3$.

Combining the matching relations (2.21) and (2.22) shows that the limits of $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ as $\boldsymbol{x}\to\boldsymbol{y}\in\mathbb{S}$ on the two sides are given in terms of the density ν_o and of the value of $\mathcal{W}_{\nu_o}$ on $\mathbb{S}$ by

$$\mathcal{W}_+(\boldsymbol{y}) = \mathcal{W}_{\nu_o}(\boldsymbol{y}) + \frac{1}{2}\nu_o \ , \quad \mathcal{W}_-(\mathbf{y}) = \mathcal{W}_{\nu_o}(\boldsymbol{y}) - \frac{1}{2}\nu_o \qquad (\forall\mathbf{y}\in\mathbb{S}) \tag{2.24}$$

Theorem 2.1.4 *[2]. Suppose Ω is a normal s.s.c. domain and $\mathbb{S}=\partial\Omega$ its closed connected boundary. Then*

$$\mathcal{W}_1(\boldsymbol{x}) = -\frac{\varpi(\boldsymbol{x})}{4\pi}, \quad \varpi(\boldsymbol{x}) := \begin{cases} 4\pi & \boldsymbol{x}\in\Omega \\ 2\pi & \boldsymbol{x}\in\partial\Omega \\ 0 & \boldsymbol{x}\in\Omega' := \mathbb{R}^3\backslash\overline{\Omega} \end{cases} \tag{2.25}$$

Hence

$$\boldsymbol{E}(\boldsymbol{x}) := -grad\,\mathcal{W}_1(\boldsymbol{x}) \equiv \boldsymbol{0} \qquad \forall\boldsymbol{x}\in\mathbb{R}^3\backslash\mathbb{S}$$

and

$$\frac{\partial \mathcal{W}_1(\boldsymbol{y})}{\partial n_y} \equiv 0 \qquad \forall \boldsymbol{y} \in \partial\Omega \tag{2.26}$$

where the normal derivative is defined on $\partial\Omega$ *by continuity as in Theorem 2.1.3* (vi).

It follows that

$$\mathcal{W}_{\nu_o}(\boldsymbol{x}) = \begin{cases} -\nu_o & \boldsymbol{x} \in \Omega \\ -\frac{1}{2}\nu_o & \mathbf{x} \in \partial\Omega \\ 0 & \boldsymbol{x} \in \Omega' := \mathbb{R}^3 \backslash \overline{\Omega} \end{cases} \tag{2.27}$$

and by removing the singularity over $\partial\Omega$ as in Theorem 2.1.3(vi) the electric field due to a double layer potential of constant density over $\partial\Omega$ is identically zero in $\mathbb{R}^3$:

$$\boldsymbol{E} = -grad\, \mathcal{W}_{\nu_o}(\boldsymbol{x}) \equiv \mathbf{0} \qquad \forall \boldsymbol{x} \in \mathbb{R}^3 \tag{2.28}$$

The double layer potential has an interesting geometrical interpretation, which goes under the name of Gauss' solid angle formula [2], inasmuch as $\varpi(\boldsymbol{x})$ in eq. (2.25) represents the total solid angle with sign, subtended by the closed connected surface $\mathbb{S} = \partial\Omega$ at the point $\boldsymbol{x}$. Indeed, if we let

$$\varphi(\boldsymbol{x}, \boldsymbol{y}) := \arccos \frac{\boldsymbol{n}(\boldsymbol{y}) \cdot (\boldsymbol{y} - \boldsymbol{x})}{|\ \boldsymbol{y} - \boldsymbol{x}|}\ , \qquad 0 \leq \varphi \leq \pi$$

denote the angle between $\boldsymbol{n}(\boldsymbol{y})$ and the vector $\boldsymbol{y}-\boldsymbol{x}$ for a generic connected surface $\mathbb{S}$ (open or closed), we have that the kernel of $-\mathcal{W}_1$

$$-\frac{\partial}{\partial n_y}\left(\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|}\right) dS_y \equiv -\frac{\boldsymbol{n}(\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})}{r^3}\, dS_y = \frac{\cos\varphi}{r^2} dS_y := d\varpi(\boldsymbol{x})$$

represents the elementary solid angle subtended by dS_y at the point $\boldsymbol{x}$, with a positive sign if $0 \leq \varphi < \pi/2$, a negative sign if $\pi/2 < \varphi \leq \pi$ (see Fig. 2.1).

Integrating over $\mathbb{S}$ we obtain the Gauss solid angle formula

$$\mathcal{W}_{\nu_o}(\mathbf{x}) = -\nu_o \frac{\varpi(\boldsymbol{x})}{4\pi} \qquad (\boldsymbol{x} \in \mathbb{R}^3) \tag{2.29}$$

Figure 2.1: Gauss' solid angle formula

which implies Theorem 2.1.4 in the case of a closed connected surface $\mathbb{S} = \partial\Omega$ (Exercise 9). If instead $\mathbb{S}$ is an open connected flat surface, a portion of a plane, the Gauss solid angle formula yields

$$(\mathcal{W}_1)_- = -\frac{1}{2}, \;\; (\mathcal{W}_1)_+ = \frac{1}{2} \qquad \Rightarrow \qquad [\mathcal{W}_1]_{\mathbb{S}} = 1$$

and

$$\mathcal{W}_1(\boldsymbol{y}) \;=\; \frac{1}{2}\;((\mathcal{W}_1)_+ + (\mathcal{W}_1)_-) = 0 \qquad \forall \boldsymbol{y} \in \mathbb{S} \tag{2.30}$$

(cfr. eq. (2.22) and Exercises 8, 10).

With this geometrical interpretation in mind, Theorem 2.1.4 can be easily extended to non-connected surfaces $\mathbb{S} = \partial\Omega$, with Ω a s.m.c. bounded domain of $\mathbb{R}^3$ (Exercise 11). The electrostatic field $\boldsymbol{E}(\boldsymbol{x}) = -grad\,\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is irrotational and solenoidal for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \mathbb{S}$, and its behavior as $\boldsymbol{x}$ approaches $\mathbb{S}$ depends on whether $\mathbb{S}$ is open or closed.

Theorem 2.1.5 (*i*) *If the surface $\mathbb{S}$ is closed the singularity of the electrostatic field $\boldsymbol{E}(\boldsymbol{x}) = -grad\,\mathcal{W}_{\nu_o}(\boldsymbol{x})$ on $\mathbb{S}$ can be removed, and $\boldsymbol{E}(\boldsymbol{x}) \equiv \boldsymbol{0}$ for $\boldsymbol{x} \in \mathbb{R}^3$.*

(*ii*) *If $\mathbb{S}$ is open, $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is discontinuous but finite as $\boldsymbol{x}$ approaches $\partial\mathbb{S}$, whereas $\boldsymbol{E}(\boldsymbol{x})$ has an unbounded singularity of the Biot-Savart type on $\partial\mathbb{S}$.*

Proof. In the case of a closed surface $\mathbb{S}$, the extended Theorem 2.1.4 (Exercise 11) shows that

$$(grad\,\mathcal{W}_{\nu_o})_+ = (grad\,\mathcal{W}_{\nu_o})_- \;=\; \boldsymbol{0}$$

on $\mathbb{S}$, and assertion (i) follows by defining $grad\,\mathcal{W}_{\nu_o}$ on $\mathbb{S}$ by continuity as in Theorem 2.1.3(vi). If $\mathbb{S}$ is open we have

$$\begin{aligned}\int_{\mathbb{S}} \frac{\partial}{\partial n_y}\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) dS_y &= -\int_{\mathbb{S}} (\boldsymbol{n}(\boldsymbol{y})-\boldsymbol{n}(\boldsymbol{x})+\boldsymbol{n}(\boldsymbol{x}))\cdot grad_x\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) dS_y \\ &= \int_{\mathbb{S}} \frac{(\boldsymbol{n}(\boldsymbol{y})-\boldsymbol{n}(\boldsymbol{x}))\cdot(\boldsymbol{x}-\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3}\, dS_y - \boldsymbol{n}(\boldsymbol{x})\cdot grad_x \int_{\mathbb{S}} \frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|}\end{aligned}$$

and as $\boldsymbol{x}\to\boldsymbol{x}_o \in \partial\mathbb{S}$

$$\frac{(\boldsymbol{n}(\boldsymbol{y})-\boldsymbol{n}(\boldsymbol{x}))\cdot(\boldsymbol{x}-\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3} = O\left(\frac{1}{|\boldsymbol{x}_o-\boldsymbol{x}|}\right)$$

so that the first integral remains bounded. As to the second integral, Exercise 6 implies that

$$\boldsymbol{n}(\boldsymbol{x})\cdot grad_x \int_{\mathbb{S}} \frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|} = O(1) \quad as \quad \boldsymbol{x}\to\boldsymbol{x}_o \in \partial\mathbb{S}$$

Thus $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ remains finite as $\boldsymbol{x}\to\boldsymbol{x}_o \in \partial\mathbb{S}$. In contrast, the gradient of $\mathcal{W}_{\nu_o}(\boldsymbol{x})$ is singular and behaves like the Biot-Savart magnetic field (1.19) in Chapter 1: the proof of this fact is deferred to a later Chapter.

We remark that the gradient of a double layer potential $\mathcal{W}_\nu$ with smooth variable density $\nu(\boldsymbol{y})$ satisfies the jump relations across $\mathbb{S}$ [31]

$$\begin{cases} \boldsymbol{n}(y)\cdot[grad\,\mathcal{W}_\nu]_{\mathbb{S}} \equiv [\frac{\partial \mathcal{W}_\nu}{\partial n}]_{\mathbb{S}} = 0 \\ \boldsymbol{n}(\boldsymbol{y})\wedge[grad\,\mathcal{W}_\nu]_{\mathbb{S}} = \boldsymbol{n}(\boldsymbol{y})\wedge grad\,\nu(\boldsymbol{y}) \end{cases} \qquad \forall \boldsymbol{y}\in\mathbb{S}$$

Thus the normal derivative of $\mathcal{W}_\nu$ can always be defined on $\mathbb{S}$ by continuity whereas this is not true for tangential derivatives unless $\nu(\boldsymbol{y})$ is constant on $\mathbb{S}$.

2.1.3 Green's identities.

The single layer, double layer and volume potentials also play a crucial role in a purely mathematical context, as a means of representing an arbitrary function. This follows from the well-known Green's (third) identity, one form of which has already been used in §1.6.1.

Theorem 2.1.6 *(Green's identity for Ω and for $\Omega' := \mathbb{R}^3 \backslash \overline{\Omega}$). (i) Let $u(\boldsymbol{x})$ denote a biregular function in a normal domain Ω of $\mathbb{R}^3$ with outer normal $\boldsymbol{n}$. Then the representation formula*

$$(2.31) \qquad \frac{1}{4\pi}\varpi(\boldsymbol{x})u(\boldsymbol{x}) \equiv \mathcal{V}_{\partial u/\partial n_-}(\boldsymbol{x}) - \mathcal{W}_{u_-}(\boldsymbol{x}) - \mathbb{V}_{\Delta_3 u}(\boldsymbol{x})$$

holds identically for $\boldsymbol{x} \in \mathbb{R}^3$, if $\varpi(\boldsymbol{x})$ is the solid angle defined in eq. (2.25) [5], *and u_- , $\partial u/\partial n_-$ are the interior traces on $\partial\Omega$ (i.e. "coming from Ω").*

(ii) Let $u(\boldsymbol{x})$ denote a biregular harmonic function in $\Omega' := \mathbb{R}^3 \backslash \overline{\Omega}$, regular at infinity and $\boldsymbol{n}$ oriented as before. Then

$$(2.32) \quad (1 - \frac{1}{4\pi}\varpi(\boldsymbol{x}))\, u(\boldsymbol{x}) \equiv -\mathcal{V}_{\partial u/\partial n_+}(\boldsymbol{x}) + \mathcal{W}_{u_+}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \mathbb{R}^3)$$

where u_+ , $\partial u/\partial n_+$ are the exterior traces on $\partial\Omega$ (i.e. "coming from Ω').

The proof, based on Green's first and second identities (Exercise 12), can be found in any textbook (see e.g. [2]).

Corollary 2.1.7 *(Green's identity for $\Omega \cup \Omega'$). Let*

$$[u] \equiv [u]_{\partial\Omega} := u_+ - u_- \quad , \quad [\partial u/\partial n] \equiv [\partial u/\partial n]_{\partial\Omega} := \partial u/\partial n_+ - \partial u/\partial n_-$$

Any function $u(\boldsymbol{x})$, biregular in Ω and in Ω', regular at infinity, and whose Laplacian $\Delta_3 u$ has compact support $K \subset \Omega$, satisfies the identity

$$(2.33) \qquad u(\boldsymbol{x}) \equiv -\mathcal{V}_{[\partial u/\partial n]}(\boldsymbol{x}) + \mathcal{W}_{[u]}(\boldsymbol{x}) - \mathbb{V}_{\Delta_3 u}(\boldsymbol{x})$$

for $\boldsymbol{x} \in \Omega \cup \Omega'$, and

$$(2.34) \qquad \frac{1}{2}\{u_-(\boldsymbol{x}) + u_+(\boldsymbol{x})\} = -\mathcal{V}_{[\partial u/\partial n]}(\boldsymbol{x}) + \mathcal{W}_{[u]}(\boldsymbol{x}) - \mathbb{V}_{\Delta_3 u}(\boldsymbol{x})$$

for $\boldsymbol{x} \in \partial\Omega$.

[5] if $\mathbf{x}_o$ is a conical point on $\partial\Omega$, eq. (31) holds with the appropriate value of the solid angle at $\mathbf{x}_o$

Stated succinctly, any biregular function is the sum of a single layer potential, a double layer potential and, if it is not harmonic, of a volume potential. When applied to an electrostatic potential, Green's identity (2.33) shows that the bounded solution of the summation problem for $\Omega \cup \Omega'$

$$\begin{cases} \Delta_3 u = -\rho(\boldsymbol{x})/\epsilon & (\boldsymbol{x} \in \Omega \cup \Omega') \\ [u(\boldsymbol{x})]_{\partial\Omega} = \nu_o \ , \quad [\frac{\partial u(\boldsymbol{x})}{\partial n}]_{\partial\Omega} = -\frac{\sigma(\mathbf{x})}{\epsilon} & (\boldsymbol{x} \in \partial\Omega) \end{cases} \tag{2.35}$$

with u regular at infinity, is represented by the formula

$$u = \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x}) + \mathcal{W}_{\nu_o}(\boldsymbol{x}) + \mathbb{V}_{\rho/\epsilon}(\boldsymbol{x}) \tag{2.36}$$

and, taking into account Theorem 2.1.5(i), the corresponding electric field is given by

$$\boldsymbol{E} = -grad\,\mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x}) - grad\,\mathbb{V}_{\rho/\epsilon}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \Omega \cup \Omega')$$

Thus the electrostatic potential u, bounded and regular at infinity, is known in $\mathbb{R}^3$ if ρ, $[u]$ and σ are known, and is identically zero if $\rho = [u] = \sigma = 0$; the bounded electrostatic field $\boldsymbol{E} = -grad\,u$ is identically zero in $\Omega \cup \Omega'$ if $\rho = \sigma = 0$, that is, in the absence of all charges. It follows that the electrostatic potential due to a charge distributed over a smooth arc of curve Γ with linear charge density χ

$$u(\boldsymbol{x}) := \int_\Gamma \frac{\chi(\boldsymbol{y})\,ds_y}{4\pi\epsilon r} \quad , \qquad r = |\boldsymbol{x} - \boldsymbol{y}| \tag{2.37}$$

which is clearly harmonic for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \Gamma$, must be singular as $\boldsymbol{x} \to \Gamma$. Indeed, it is easy to see that

$$\lim_{\boldsymbol{x}\to\Gamma} |u(\boldsymbol{x})| = +\infty \quad , \qquad \lim_{\boldsymbol{x}\to\Gamma} |grad\,u(\boldsymbol{x})| = +\infty$$

and the singularity of $u(\boldsymbol{x})$ turns out to be weaker than the Coulomb singularity (2.10) or the dipole singularity (2.11) (Exercise 13).

Eq. (2.33) also shows that any biregular harmonic function u in $\Omega \cup \Omega'$, regular at infinity, is represented by a superposition of a single layer potential with density $-[\frac{\partial u}{\partial n}]_{\partial\Omega}$ and a double layer potential with density $[u(\boldsymbol{x})]_{\partial\Omega}$:

$$u(\boldsymbol{x}) \equiv -\mathcal{V}_{[\partial u/\partial n]}(\boldsymbol{x}) + \mathcal{W}_{[u]}(\mathbf{x}) \qquad (\boldsymbol{x} \in \Omega \cup \Omega') \tag{2.38}$$

If u is continuous across $\partial\Omega$, then $[u] = 0$ and u can be represented as a single layer

$$u(\boldsymbol{x}) \equiv -\mathcal{V}_{[\partial u/\partial n]}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \Omega \cup \Omega') \tag{2.39}$$

if u has continuous normal derivative across $\partial\Omega$, then $[\partial u/\partial n] = 0$ and u can be represented as a double layer

$$u(\mathbf{x}) \equiv \mathcal{W}_{[u]}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \Omega \cup \Omega') \tag{2.40}$$

It follows that $u \equiv 0$ if $[u]_{\partial\Omega} = [\partial u/\partial n]_{\partial\Omega} = 0$:any biregular harmonic function u in $\Omega \cup \Omega'$ regular at infinity and continuous together with its normal derivative across $\partial\Omega$ is identically zero.

2.2 Single conductor in a homogeneous dielectric

We consider here electrostatic problems which arise from assigning the value of the potential V or of the total surface charge Q on the boundary of a single homogeneous bounded conductor C surrounded by a homogeneous unbounded dielectric $D = \mathbb{R}^3 \backslash \overline{C}$ having permittivity ϵ.

2.2.1 Potential and charge problem.

We will assume that $\rho(\boldsymbol{x}) \equiv 0$ everywhere and that the potential u is continuous in D and regular at infinity. These conditions imply that the plane at infinity is grounded, i.e. is kept at the potential $u_\infty = 0$, and exclude the presence of point charges and dipoles.

A. Potential problem: Find the potential u in D, regular at infinity, and the induced surface charge density σ on ∂C if the latter is a connected bounded surface having an assigned potential V.

The potential u satisfies the exterior Dirichlet problem for the domain $D = \mathbb{R}^3 \backslash \overline{C}$ with a constant boundary value V

$$\begin{aligned} &\Delta_3\, u = 0 && \text{in } D = \mathbb{R}^3 \backslash \overline{C} \\ &u(\boldsymbol{y}) = V && \boldsymbol{y} \in \partial C \\ &u(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-1}) && \text{uniformly as } |\boldsymbol{x}| \to \infty \end{aligned} \tag{2.41}$$

where the boundary condition on $\partial D = \partial C$ is to be interpreted in the sense of the limit from the exterior. The unique solution [6] $u(\boldsymbol{x})$ of (2.41) has the form(Exercise 14)

$$u = V u_1 \tag{2.42}$$

where $u_1(\boldsymbol{x})$ is the *capacitary potential* of ∂C, i.e. the unique solution of the charge problem (2.41) for a unit tension $V = 1$. The capacitary potential is then the solution of the exterior Dirichlet problem

$$\begin{aligned} \Delta_3\, u_1 &= 0 && \text{in } \; D = \mathbb{R}^3 \backslash \overline{C} \\ u_1(\boldsymbol{y}) &= 1 && \mathbf{y} \in \partial C \\ u_1(\boldsymbol{x}) &= O(\,|x|^{-1}) && \text{uniformly as } |\boldsymbol{x}| \to \infty \end{aligned} \tag{2.43}$$

It is well-known that u_1 exists, is unique, is positive and is biregular in D if ∂C is of class C^2 [2]. The tension V induces a surface charge distribution on ∂C with density σ given by

$$\sigma(\boldsymbol{y}) := -\epsilon \frac{\partial u(\boldsymbol{y})}{\partial n} = -\epsilon V \frac{\partial u_1(\boldsymbol{y})}{\partial n}\,, \; \boldsymbol{y} \in \partial C \tag{2.44}$$

where $\boldsymbol{n}$ is the normal to ∂C oriented towards the interior of D at the point $\boldsymbol{y}$, and the normal derivative $\partial u(\boldsymbol{y})/\partial n$ is to be interpreted in the sense of eq. (2.4). The total surface charge follows by integrating $\sigma = V\sigma_1$ over ∂C

$$Q := \int_{\partial C} \sigma dS = -\epsilon V \int_{\partial C} \frac{\partial u_1}{\partial n}\, dS \tag{2.45}$$

Definition 2.2.1 *The capacity of the connected conducting surface ∂C is the ratio*

$$\mathbb{C} := \frac{Q}{V} = \int_{\partial C} \sigma_1\, dS \tag{2.46}$$

where

$$\sigma_1 := -\epsilon \frac{\partial u_1}{\partial n} \equiv V^{-1}\sigma$$

is the surface charge density induced by the capacitary potential.

[6] uniqueness can be proven using the maximum principle [2] or the first Green identity (E3) of Exercise 12

We may also refer the capacity $\mathbb{C}$ of the closed connected surface ∂C to the domain C bounded by the surface. In this way the capacity becomes a set function defined over a suitable collection of sets in $\mathbb{R}^3$ [2].

Proposition 2.2.2 *The capacitary potential coincides with the single layer potential*

$$u_1 = \mathcal{V}_{\sigma_1/\epsilon}(\boldsymbol{x}) \tag{2.47}$$

with density σ_1 determined as the unique solution of the boundary integral equation of the first kind on ∂C

$$\mathcal{V}_{\sigma_1/\epsilon}(\boldsymbol{x}) = 1 \ , \ \ \boldsymbol{x} \in \partial C \tag{2.48}$$

Proof. Eq. (2.48) has a unique continuous solution σ_1 [34] and $\mathcal{V}_{\sigma_1/\epsilon}(\boldsymbol{x})$ has all the required properties by force of Theorem 2.1.2. In particular the uniqueness theorem for the interior Dirichlet problem [2] implies that $\mathcal{V}_{\sigma_1/\epsilon}(\boldsymbol{x}) \equiv 1$ for all $\boldsymbol{x} \in C$, and the capacitary potential is thus extended inside the conductor with the constant value $u_1 = 1$.

Corollary 2.2.3 *The unique solution of the Dirichlet problem (2.41) under the stated assumptions is*

$$u = V\mathcal{V}_{\sigma_1/\epsilon}(\boldsymbol{x}) \equiv \mathcal{V}_{\sigma\,/\epsilon}(\boldsymbol{x})$$

where the charge density $\sigma \in C^o(\partial C)$ is the unique solution of the boundary integral equation of the first kind

$$\mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x}) = V \ \ , \ \ \boldsymbol{x} \in \partial C \tag{2.49}$$

Proposition 2.2.4 *The capacity $\mathbb{C}$ is positive and can be interpreted as (twice the value of) the energy stored in the dielectric external to the conductor C when the conductor has unit potential $V = 1$.*

Proof. The electrostatic energy $\mathcal{E} = \mathcal{E}_D$ is given by

$$\mathcal{E} = \frac{1}{2}\epsilon \int_D |\boldsymbol{E}(\boldsymbol{x})|^2 \, dx \tag{2.50}$$

Since $\boldsymbol{E}= -grad\, u$, $\mathcal{E}$ is given by the Dirichlet integral of u [2] times $\epsilon/2$

$$\mathcal{E} = \frac{1}{2}\epsilon \int_D |grad\, u|^2 \boldsymbol{dx}$$

But $u = V u_1$, so that

$$\mathcal{E} = \frac{1}{2}\epsilon V^2 \int_D |grad\, u_1|^2\, \boldsymbol{dx}$$

Since by definition $u_1 = 1$ on ∂C, eq.(2.45) and Green's first identity (E3) show that $\mathbb{C}$ can be written in the form

$$\mathbb{C} = -\epsilon \int_{\partial C} u_1 \frac{\partial u_1}{\partial n}\, dS = \ \epsilon \int_D |grad\, u_1|^2\, dV \tag{2.51}$$

It follows that

$$\mathcal{E} = \frac{1}{2}\mathbb{C}V^2 \equiv \frac{1}{2}Q^2/\mathbb{C} \tag{2.52}$$

and $\mathbb{C} = 2\mathcal{E}/V^2 > 0$.

Remark 2. The capacitary potential depends only on the geometry of ∂C, and so does the capacity of a surface, apart from the factor ϵ. For example, the capacity of a sphere of radius R is $\mathbb{C} = 4\pi\epsilon R$ (Exercise 15).

Remark 3. Corollary 2.2.3 and eq. (2.44) imply that the unique solution σ of eq. (2.49) satisfies the linear fixpoint equation on ∂C

$$-\frac{\partial}{\partial n}\mathcal{V}_\sigma = \sigma \tag{2.53}$$

which has the form of a homogeneous Fredholm integral equation of the second kind [34]. The solution σ of this equation is called a Robin density and the corresponding single layer potential $\mathcal{V}_\sigma$ a *Robin potential* of C [2]. **Remark 4.** As already mentioned, the solution (2.42) of the potential problem is not fully general. Experimental evidence starting from Volta's original discoveries shows that the potential u has in general a constant jump $[u] = \nu_o$ across the conductor surface ∂C. Taking this jump into account, the full solution of the potential problem would be given by the sum of a single and double layer potential

$$u \ = V u_1 + \mathcal{W}_{\nu_o} \equiv \ \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x}) + \frac{\nu_o}{4\pi}\int_{\partial C} \frac{\partial}{\partial n_y}\Big(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\Big)\, dS_y \tag{2.54}$$

However, the electric field $\boldsymbol{E}$, by virtue of Theorem 2.1.5, is not influenced by the potential jump ν_o and is still given by

$$\boldsymbol{E}(\boldsymbol{x}) = -grad\, \mathcal{V}_{\sigma/\epsilon}(\boldsymbol{x})$$

Similarly, by force of eqs. (2.26) and (2.44), the charge density σ remains unaltered and is still given by the solution of eq. (2.49) as before. In the case of N homogeneous conductors C_i, treated in detail in §2.3, the potential jumps $[u]_{\partial C_i} = \nu_i$ give rise to an additional potential of the form of the sum of double layers

$$\sum_{i=1}^{N} \mathcal{W}_{\nu_i} \equiv \sum_{i=1}^{N} \frac{\nu_i}{4\pi} \int_{\partial C_i} \frac{\partial}{\partial n_y}\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) dS_y$$

whose contributions to the electric field in D and to the surface charges on ∂C_i are still zero.

Therefore the Electrostatics of homogeneous conductors can be carried out, as we do here, neglecting the potential jumps. The situation changes in the case of non-homogeneous conductors, as e.g. in the case of the Volta cell [27].

Remark 5. In the presence of a given uniform field $\boldsymbol{E}_o$ in D, $u(\boldsymbol{x})$ is no longer regular at infinity and the asymptotic condition (2.5) is replaced with

$$\text{(2.55)}\quad \boldsymbol{E}(\boldsymbol{x}) = -grad\, u(\boldsymbol{x}) \sim \boldsymbol{E}_o, u(\boldsymbol{x}) \sim -\boldsymbol{E}_o \cdot \boldsymbol{x} + O(|\boldsymbol{x}|^{-1}) \;\; as\, |\boldsymbol{x}| \to +\infty$$

(Exercise 16).

Remark 6. For a bounded dielectric D surrounded by an unbounded conductor C the potential problem is trivial, since $\sigma \equiv 0$ by eqs. (2.6), (2.9) and the corresponding interior Dirichlet problem

$$\Delta_3\, u = 0 \quad \text{in} \;\; D\,, \qquad u = V \quad \text{on}\; \partial D$$

has the unique solution $u = V$, $\boldsymbol{E} = \boldsymbol{0}$, $Q = 0$.

B. Charge problem: Find the potential u in $\overline{\mathrm{D}}$ and the surface charge density σ, if the total surface charge Q is given on the (connected) conductor boundary ∂C.

The solution of the charge problem, with u regular at infinity, can be written in terms of the capacitary potential u_1. Let $\mathbb{C}$ be the capacity of ∂C, and let

$$V := \frac{Q}{\mathbb{C}} \tag{2.56}$$

Proposition 2.2.5 *The unique solution of the charge problem is given by*

$$u = V\mathcal{V}_{\sigma_1/\epsilon} \equiv Vu_1 \quad , \ \sigma = V\sigma_1 \equiv \frac{Q}{\mathbb{C}}\sigma_1 \tag{2.57}$$

where u_1 is the capacitary potential and σ_1 is the solution of the boundary integral equation (2.48).

Proof. The previous results, in particular eq. (2.53), and the uniqueness theorem show that (2.57) is the solution.

For a bounded dielectric surrounded by an unbounded conductor the charge problem has no meaning, since necessarily $\sigma = 0$ (see Remark 6 above).

2.2.2 Influence problems: the Green function.

Two further interesting problems in Electrostatics (or potential theory) motivate the introduction of the mathematical concept of Green function, or influence function.

C. Influence problem for a grounded conducting surface: Given a homogeneous dielectric D bounded by a grounded conducting surface ∂D, find the potential u in D and the density σ_{ind} of the induced charge on ∂D due to a point charge Q_o concentrated at an interior point $\mathbf{y} \in D$.

We are assuming that D is a normal domain with a C^2 boundary ∂D, that $u = 0$ on ∂D, and that $D' = \mathbb{R}^3 \backslash \overline{D}$ is another dielectric medium, whose physical properties as we will see are entirely irrelevant (see Remark 7 below). It is clear that the solution $u(\boldsymbol{x})$ will depend also on the position $\boldsymbol{y}$ of the influencing charge Q_o. If we let

$$u(\boldsymbol{x}) = \frac{Q_o}{\epsilon} G(\boldsymbol{x}|\boldsymbol{y}) \tag{2.58}$$

for $\boldsymbol{x}{\neq}\boldsymbol{y}$, we expect that as $\boldsymbol{x}$ approaches $\boldsymbol{y}$ the function $G(\boldsymbol{x}|\boldsymbol{y})$, called Green's function of the Laplacian in D, has the singular behavior $\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}$ appropriate to the Coulomb potential due to the point charge Q_o. Moreover $G(\boldsymbol{x}|\boldsymbol{y})$ must vanish for $\boldsymbol{x}\in \partial D$. The Green function $G(\boldsymbol{x}|\boldsymbol{y})$ must therefore be a solution of the Dirichlet problem in the punctured domain $D\backslash\{\boldsymbol{x}\}$

$$
\begin{aligned}
&\Delta_3^x G(\boldsymbol{x}|\boldsymbol{y}) \equiv \sum_{k=1}^{3}\frac{\partial^2 G(\boldsymbol{x}|\boldsymbol{y})}{\partial x_k^2} = 0 \qquad \text{for } \boldsymbol{x}\in D\backslash\{\boldsymbol{y}\} \\
&G(\boldsymbol{x}|\boldsymbol{y}) \sim \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} \qquad as\ |\boldsymbol{x}-\boldsymbol{y}|\to 0 \\
&G(\boldsymbol{x}|\boldsymbol{y}) = 0 \qquad for\ \boldsymbol{x}\in\partial D
\end{aligned}
\tag{2.59}
$$

for every $\boldsymbol{y}\in D$. If we set

$$
G(\boldsymbol{x}|\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} + g(\boldsymbol{x}|\boldsymbol{y}) \tag{2.60}
$$

the auxiliary function $g(\boldsymbol{x}|\boldsymbol{y})$ has a removable singularity at the point $\boldsymbol{x}{=}\boldsymbol{y}$ and is a solution of the ordinary interior Dirichlet problem

$$
\begin{cases}
\Delta_3^x g \equiv \sum_{k=1}^{3}\frac{\partial^2 g(\boldsymbol{x}|\boldsymbol{y})}{\partial x_k^2} = 0 & for \boldsymbol{x}\in D \\
g(\boldsymbol{x}|\boldsymbol{y}) = -\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} & for \boldsymbol{x}\in\partial D
\end{cases}
\tag{2.61}
$$

for every $\boldsymbol{y}\in D$. As ∂D is smooth by assumption and the boundary data $(-4\pi|\boldsymbol{x}-\boldsymbol{y}|)^{-1}$ are of class C^∞ for all $\boldsymbol{y}\in D$, the unique solution g of (2.61) is a biregular harmonic function of $\boldsymbol{x}$ for all $\boldsymbol{y}\in D$. Hence the Green function $G(\boldsymbol{x}|\boldsymbol{y})$ exists, is unique, and has continuous normal derivative $\partial G/\partial n$ on ∂D [2].

Proposition 2.2.6 . *(i) $G(\boldsymbol{x}|\boldsymbol{y})$ is symmetric, $G(\boldsymbol{x}|\boldsymbol{y}) = G(\boldsymbol{y}|\boldsymbol{x})$ $\forall\boldsymbol{x}\in D$, $\boldsymbol{y}\in D$*

(ii) $G(\boldsymbol{x}|\boldsymbol{y}) > 0$ for every $\boldsymbol{x}\in D$, $\boldsymbol{y}\in D$

(iii) If $\boldsymbol{n}_x$ is the interior normal at the point $\boldsymbol{x}\in\partial D$ oriented towards the interior of D, then

$$
\frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} \geq 0 \ , \quad \int_{\partial D}\frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x}\, dS_x \equiv 1 \qquad \forall\boldsymbol{y}\in D \tag{2.62}
$$

Proof. See [2] and Exercise 17.

The corresponding electrostatic potential u, given by eq. (2.58), gives rise to an induced surface charge density at a point $\boldsymbol{x}\in \partial D$ given by

$$\sigma_{ind}(\boldsymbol{x}|\boldsymbol{y}) = -\epsilon \frac{\partial u}{\partial n_x} = -Q_{\text{o}} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} \tag{2.63}$$

so that, by force of eq. (2.62), the total induced charge on ∂D is

$$Q_{ind} = -Q_{\text{o}} \int_{\partial D} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} \, dS_x \equiv -Q_o$$

for all $\boldsymbol{y} \in D$. In other words, the induced charge Q_{ind} is equal and opposite to the influencing charge Q_o whatever the position of the latter in D, in accordance with the Gauss Law.

The physical interpretation of the Green function is clear: $Q_{\text{o}}G(\boldsymbol{x}|\boldsymbol{y})/\epsilon$ represents the electrostatic potential at the point $\boldsymbol{x}$, due to the influencing charge at the point $\boldsymbol{y}$, and is the sum of the Coulomb potential $\frac{Q_o}{4\pi\epsilon|\boldsymbol{x}-\boldsymbol{y}|}$ due to Q_o and of the potential $Q_{\text{o}}g(\boldsymbol{x}|\boldsymbol{y})/\epsilon$ due to the distribution of the induced charges on ∂D with density σ_{ind}. These induced surface charges guarantee that the grounded conducting surface ∂D has zero potential. The symmetry of the Green function (see Proposition 2.2.6(i)) implies a reciprocity relation which says that the potential at $\boldsymbol{x}$ due to the influencing charge in $\boldsymbol{y}$ is the same as the potential at $\boldsymbol{y}$ due to the influencing charge concentrated in $\boldsymbol{x}$.

Remark 7. The potential and the electrostatic field vanish identically in the exterior dielectric $D' = \mathbb{R}^3 \backslash \overline{D}$.

Remark 8. The Green function G depends solely on the domain D and can be determined explicitly in the case of symmetrical domains like e.g. a sphere, an ellipsoid, or a half-space [7], using the method of images. The idea underlying this method is to replace the potential of the induced charges $g(\boldsymbol{x}|\boldsymbol{y})$ with the Coulomb potential due to a virtual point charge Q'_o concentrated at a point $\boldsymbol{y}' \in \mathbb{R}^3 \backslash \overline{D}$, with Q'_o and $\boldsymbol{y}'$ chosen in such a way that $G(\boldsymbol{x}|\boldsymbol{y}) = 0$ for every $\boldsymbol{x} \in \partial D$ ((Exercise 18).

A related influence problem occurs when the (connected) conducting surface ∂D, instead of being grounded, is isolated, so that its total induced surface charge $Q_{ind} = 0$ on ∂D. Its constant potential V is unknown and

[7] in the case of a half-space the domain D is unbounded, see Exercise 18

must be determined from this condition: the result is $V = Q_o/\mathbb{C}$ if $\mathbb{C}$ is the capacity of ∂D (Exercise 19). More generally, we can consider the following problem (see Fig. 2.2).

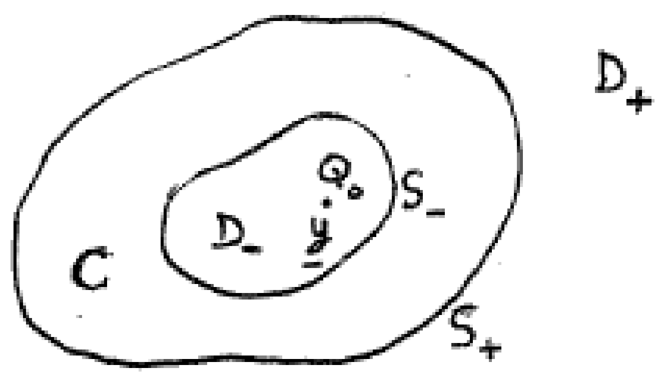

Figure 2.2: Isolated conducting sheath

D. Influence problem for an isolated conducting sheath: Given an inner dielectric D_- and an outer dielectric D_+ separated by an isolated conducting sheath C bounded by two closed connected surfaces S_- and S_+, find the potential u, regular at infinity, due to an influencing point charge Q_o concentrated at a point $\mathbf{y} \in D_-$. We are assuming that D_- has permittivity ϵ_-, $D_+ = \mathbb{R}^3 \setminus \overline{D_- \cup C}$ has permittivity ϵ_+, and $\partial D_\mp = S_\mp$. The electrostatic potential is given by

$$u(\boldsymbol{x}) = \begin{cases} V u_1(\boldsymbol{x}) & \text{in } D_+ \\ V & \text{in } C \\ \frac{Q_o}{\epsilon_-} G(\boldsymbol{x}|\boldsymbol{y}) + V & \text{in } D_- \end{cases}$$

where $u_1(\boldsymbol{x})$ is the capacitary potential of S_+, $G(\boldsymbol{x}|\boldsymbol{y})$ is the Green function of D_-, and V is unknown. The surface charges induced over S_+ and S_- have densities given by

$$\sigma_+(\boldsymbol{x}|\boldsymbol{y}) = -\epsilon_+ V \frac{\partial u_1}{\partial n_x} \qquad , \qquad \sigma_-(\boldsymbol{x}|\boldsymbol{y}) = -Q_o \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x}$$

where $\boldsymbol{n}_x$ is the unit normal on S_+ and S_- oriented towards the exterior of

C. By force of eq. (2.62) we have

$$\int_{S_-} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} dS_x \equiv 1 \qquad \forall\, \boldsymbol{y} \in D$$

and so the total induced charges are

$$Q_+ = V\mathbb{C}_+ \qquad , \qquad Q_- := -Q_o \int_{\partial D} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} dS_x = -Q_o$$

where $\mathbb{C}_+$ is the capacity of the outer surface S_+ (cfr. Definition 1). Since C is isolated we have $Q_{ind} = Q_- + Q_+ = 0$, and therefore

$$Q_+ = Q_o$$

The potential V of C is then given by

$$V = \frac{Q_o}{\mathbb{C}_+}$$

(Exercise 20). These results are confirmed by Faraday's celebrated "ice-pail experiment": a point charge Q_o placed at the interior of the ice-pail gives rise to an equal total charge $Q_+ = V\mathbb{C}_+$ on the exterior metal surface that can be revealed by an electrometer reading.

2.3 The fundamental problems of Electrostatics

2.3.1 Potential and charge problems for N conductors.

Problems A and B of the previous section can be generalized to a set of N homogeneous bounded conductors C_i surrounded by an uncharged dielectric D with constant permittivity ϵ. To simplify the exposition, we will suppose that all the boundary surfaces ∂C_i of the conductors are connected (i.e. that the C_i's are s.s.c.).

Problem I: Given the N constant values V_i of u on ∂C_i, find the potential u, harmonic in the exterior domain D and regular at infinity, and the surface charges on ∂C_i ($i = 1, ..., N$).

Let us introduce the capacitary potentials u_i of ∂C_i, defined as the unique solutions, biregular in D, of the N exterior Dirichlet problems for $i = 1, ..., N$:

$$\begin{aligned} &\Delta_3 \, u_i = 0 && \text{in } \; D = \mathbb{R}^3 \setminus \bigcup_{i=1}^{N} \overline{C}_i \\ &u_i = \delta_{ij} && \text{on } \partial C_j \; , \; j = 1, ..., N \\ &u_i \; \text{ regular at infinity} \end{aligned} \tag{2.64}$$

where δ_{ij} is the Kronecker delta. The surface charge densities $\sigma_{i,j}$ induced by u_i on the conductor surface ∂C_j is

$$\sigma_{i,j} := -\epsilon \, \frac{\partial u_i}{\partial n_j} \qquad (i, j = 1, ..., N) \tag{2.65}$$

where $\boldsymbol{n}_j$ is the outer normal to ∂C_j , oriented towards D. Problem I has then the unique solution, biregular in D,given by the formula

$$u(\boldsymbol{x}) = \sum_{i=1}^{N} V_i \, u_i(\boldsymbol{x}) \;\; for \, \boldsymbol{x} \in D \; , \;\; u = V_k \;\; for \;\; \boldsymbol{x} \in C_k \;\; (k = 1, ..., N) \tag{2.66}$$

which generalizes (2.42). From eqs. (2.65) and (2.66) it follows that the surface charge densities

$$\sigma_k := -\epsilon \, \frac{\partial u}{\partial n_k} \qquad (k = 1, ..., N)$$

induced by u on the conductor surface $S_k = \partial C_k$ are given by

$$\sigma_k = \sum_{i=1}^{N} V_i \, \sigma_{i,k} \qquad (k = 1, ..., N) \tag{2.67}$$

and the total surface charges are

$$Q_k = \int_{S_k} \sigma_k dS = \sum_{i=1}^{N} V_i \int_{S_k} \sigma_{i,k} dS \tag{2.68}$$

In conclusion, the solution u of problem I is given by eq. (2.66) and vanishes if and only if, all the potentials V_i are zero; the surface charges Q_k then follow from eqs. (2.65) and (2.68).

Problem II: given the N total surface charges Q_i on ∂C_i, find the potential u, harmonic in the exterior domain D and regular at infinity, and the values V_i of u on ∂C_i $(i = 1, ..., N)$.

In order to solve this problem, let us define the capacity coefficients of the set of conductors

$$\mathbb{C}_{ij} := \int_{\partial C_j} \sigma_{i,j} dS \equiv -\epsilon \int_{\partial C_j} \frac{\partial u_i}{\partial n_j} dS \qquad (i, j = 1, ..., N) \tag{2.69}$$

(cfr. eq. (2.46)) which are known (at least in line of principle) and form the entries of the $N \times N$ capacity matrix

$$\mathbb{C} = [\mathbb{C}_{ij}]$$

Eq. (2.68) can then be written in the form

$$Q_k = \sum_{i=1}^{N} \mathbb{C}_{ki} V_i \qquad (k = 1, ..., N) \tag{2.70}$$

Proposition 2.3.1 *The capacity matrix is symmetric and positive definite, and the electrostatic energy in D is given by the quadratic form*

$$\mathcal{E}_D = \frac{1}{2} \sum_{j=1}^{N} \mathbb{C}_{ij} V_i V_j$$

Proof. Taking into account the definition of the capacitary potentials, Green's second identity (E4) applied to the harmonic functions u_i and u_j in the exterior domain D yields

$$0 = \int_{\partial D} (u_i \frac{\partial u_j}{\partial n} - u_j \frac{\partial u_i}{\partial n}) \, dS = \int_{\partial C_i} \frac{\partial u_j}{\partial n_i} dS - \int_{\partial C_j} \frac{\partial u_i}{\partial n_j} \, dS$$

that is,

$$\mathbb{C}_{ij} = \mathbb{C}_{ji} \qquad (i, j = 1, ..., N)$$

By force of Green's first identity (E3) applied to the harmonic function u in the external domain D, the electrostatic energy in D

$$\mathcal{E}_D = \int_D \frac{1}{2} \epsilon |\boldsymbol{E}(\boldsymbol{x})|^2 \, \boldsymbol{dx} \equiv \frac{1}{2} \epsilon \int_D |grad\, u(\boldsymbol{x})|^2 \boldsymbol{dx}$$

takes the form

$$\mathcal{E}_D = -\frac{1}{2}\epsilon \int_{\partial D} u\frac{\partial u}{\partial n}\, dS = \frac{1}{2}\sum_{k=1}^{N}\int_{\partial C_k} u\sigma_k \, dS \tag{2.71}$$

since $\partial D = \underset{i}{\cup}\partial C_i$. Substituting the expression (2.66) for u and taking into account the boundary condition $u\big|_{\partial C_i} = V_i$ we find

$$\mathcal{E}_D = -\frac{1}{2}\epsilon \sum_{i,j,k=1}^{N} V_i V_j \int_{\partial C_k} \frac{\partial u_i}{\partial n_k}\, dS = \frac{1}{2}\sum_{j=1}^{N} \mathbb{C}_{ij} V_i V_j$$

This quadratic form is positive definite, because $\mathcal{E}_D$ is non-negative and is zero if and only if u and V vanish. Therefore the matrix $\mathbb{C}$ is positive definite and $det\,\mathbb{C} \neq 0$.

It follows that $\mathbb{C}^{-1}$ exists and the inverse matrix $\mathbb{P} = [\mathbb{P}_{ij}] := \mathbb{C}^{-1}$, called the potential matrix, is also symmetric and positive definite.

Proposition 2.3.2 *The solution of Problem II is given by*

$$u(\boldsymbol{x}) = \sum_{i,j=1}^{N} \mathbb{P}_{ij} Q_j u_i(\boldsymbol{x}) \qquad (\boldsymbol{x} \in D) \tag{2.72}$$

In other words, the solution is of the form (2.66), with the V_i's given by

$$V_i = \sum_{j=1}^{N} \mathbb{P}_{ij} Q_j \qquad (i = 1, ..., N) \tag{2.73}$$

Proof. $u(\boldsymbol{x})$ is harmonic in D and regular at infinity, and as $\boldsymbol{x} \in D$ approaches ∂C_k we have

$$u\big|_{\partial C_k} = \sum_{i,j=1}^{N} \mathbb{P}_{ij} Q_j \delta_{ik} = \sum_{j=1}^{N} \mathbb{P}_{kj} Q_j = V_k$$

By force of eqs. (2.65) and (2.69) the total surface charge density on ∂C_k is given by

$$-\epsilon \int_{\partial C_k} \frac{\partial u}{\partial n}\, dS \;= -\epsilon \sum_{i,j=1}^{N} \mathbb{P}_{ij}\; Q_j \int_{\partial C_k} \frac{\partial u_i}{\partial n_k}\, dS = \sum_{i,j=1}^{N} \mathbb{C}_{ki}\mathbb{P}_{ij}Q_j$$

and since $\mathbb{P} = \mathbb{C}^{-1}$ this is equal to Q_k, as required.

We observe that, if we introduce the N-vectors $\boldsymbol{V} = (V_1, ..., V_N)$, $\mathbf{Q} = (Q_1, ..., Q_N)$ and $\mathbf{u} = (u_1, ..., u_N)$, eqs. (2.70), (2.72) and (2.73) become

$$\boldsymbol{Q} = \mathbb{C}\,\boldsymbol{V}\ ,\ \ \boldsymbol{u} = \boldsymbol{V}\cdot\boldsymbol{u}\quad ,\quad \boldsymbol{V} \;= \mathbb{P}\boldsymbol{Q}$$

and the energy can be written in terms of the surface charges as

$$\mathcal{E}_D \;= \frac{1}{2}\boldsymbol{V}\cdot\mathbb{C}\;\boldsymbol{V} \equiv \frac{1}{2}\boldsymbol{V}\cdot\boldsymbol{Q} \equiv \frac{1}{2}\mathbb{P}\boldsymbol{Q}\cdot\boldsymbol{Q} \tag{2.74}$$

In conclusion, the solution u of problem II is given by eq. (2.72), and vanishes if, and only if, all the surface charges Q_j are zero.

Remark 9. Apart from the coefficient ϵ, the capacitary potential depends only on the geometry of the conductors ∂C_i, and hence so do the capacity matrix and the potential matrix.

2.3.2 Condensers.

Problem II is of particular interest when the system of conductors is neutralized, that is, when the sum of all the surface charges Q_i vanishes

$$\sum_{i=1}^{N} Q_i \;=\; 0$$

For $N = 2$ a set of two conductors C_1, C_2 is neutralized if the respective charges are, say,

$$Q_1 = Q \quad , \quad Q_2 = -Q \tag{2.75}$$

Such a configuration is commonly called a condenser when the two conductors are close, so that the electric field is concentrated in the space between the two condenser plates P_1 and P_2. If $V = V_1 - V_2$ is the potential difference between the two plates, we may suppose without loss of generality that $V_1 = V,\ V_2 = 0$. Eq. (2.70) takes then the form

$$\mathbb{C}_{11}V = Q,\ \ \mathbb{C}_{21}V = -Q$$

and so $\mathbb{C}_{21} = -\mathbb{C}_{11}$, $\mathbb{C}_{11} = Q/V$. The quantity $\mathbb{C} = \mathbb{C}_{11} = \mathbb{C}_{22}$, that is to say

$$\mathbb{C} := \frac{Q}{V_1 - V_2} \tag{2.76}$$

is defined as the capacity of the condenser.

For example, a plane condenser is formed by two plane parallel plates of equal area A and distance $h \ll \sqrt{A}$, say at $z = 0$, $z = h$. The electric field and the potential between the two plates are then approximately given by eq. (2.12) and (2.13):

$$\boldsymbol{E} = \frac{\sigma}{\epsilon}\boldsymbol{n}, \quad u = -\frac{\sigma}{\epsilon}(z-h) = -\frac{Q}{A\epsilon}(z-h) \qquad (0 \le z \le h)$$

where $\boldsymbol{n}$=$\mathbf{c}_3$ and $\sigma = Q/A$ is the surface charge density on the (interior face of) the first plate. It follows that $V = Qh/A\epsilon$, and the capacity of the plane condenser

$$\mathbb{C} = \frac{Q}{V} \cong \frac{A\epsilon}{h} \tag{2.77}$$

is proportional to the plate area and inversely proportional to the plates distance. In the case of a spherical condenser the plates P_1, P_2 are concentric spheres of radii $R_1 < R_2$ and the capacity $\mathbb{C} = \mathbb{C}_{11}$ is given exactly by

$$\mathbb{C} = \frac{4\pi\epsilon R_1 R_2}{R_2 - R_1} \tag{2.78}$$

(Exercise 21). Letting $R_2 \to +\infty$ we re-obtain the expression in Remark 2 for the capacity of an isolated conducting sphere of radius $R = R_1$. The electric field in a spherical condenser furnishes a simple example of a solenoidal vector field having no global vector potential (Exercise 22).

The relation between condenser tension and charge

$$V(t) = \frac{Q(t)}{\mathbb{C}}$$

remains valid in the case of quasi-stationary fields.

2.4 Kelvin's and Earnshaw's theorems.

Suppose that the system of rigid conducting bodies C_i $(i = 1, ..., N)$ embedded in the uncharged dielectric D, considered in §2.3.1, has d degrees of freedom, so that its position is identified by d Lagrangian coordinates $q_1, ..., q_d$ $(d = 6N$ in the absence of mechanical constraints). We have seen in eq. (2.74) that the energy of the system of conductors is given by the quadratic form

$$\mathcal{E} = \frac{1}{2}\mathbb{P}\boldsymbol{Q}\cdot\boldsymbol{Q} \equiv \frac{1}{2}\sum_{j=1}^{N} \mathbb{P}_{ij}Q_iQ_j$$

where $\mathbb{P}$ is the potential matrix and Q_j is the total electric surface charge of the $j-$th conductor. Since $\mathbb{P}_{ij}$ (and Q_i in case of Problem I) depend on the Lagrangian coordinates, the energy will also be a function of $q_1, ..., q_d$:

$$\mathcal{E} = \mathcal{E}(q_1, ..., q_d)$$

The presence of the charges $Q_1, ..., Q_N$ gives rise to a system of Lagrangian forces $\mathcal{Q}_1, ..., \mathcal{Q}_d$, called ponderomotive forces, acting on the rigid conductors [8]. These forces can be best calculated in terms of derivatives of the energy. The work

$$dL = \sum_{k=1}^{d} \mathcal{Q}_k dq_k$$

done by $\mathcal{Q}_1, ..., \mathcal{Q}_d$ in correspondence to an infinitesimal variation of the Lagrangian coordinates $dq_1, ..., dq_d$ must balance the variation of energy $d\mathcal{E}$:

$$d\mathcal{E} + dL = 0$$

Since

$$d\mathcal{E} = \sum_{k=1}^{d} \frac{\partial \mathcal{E}}{\partial q_k} dq_k$$

we obtain

$$\mathcal{Q}_k(q_1, ..., q_d) = -\frac{\partial \mathcal{E}}{\partial q_k} \qquad (k = 1, ..., d)$$

[8] We adopt the usual notation for the Lagrangian coordinates and forces, which should not be confused with electric charges. See e.g. H. Goldstein, Classical Mechanics, J. Wiley & Sons.

The system of conductors will be in mechanical equilibrium in a certain configuration $\overline{q}_1, ..., \overline{q}_d$ if

$$\mathcal{Q}_k(\overline{q}_1, ..., \overline{q}_d) = -\frac{\partial \mathcal{E}}{\partial q_k}\Big|_{q=\overline{q}} = 0 \qquad \text{for all } k = 1, ..., d$$

that is, if $\mathcal{E}$ is stationary at $\overline{q} = (\overline{q}_1, ..., \overline{q}_d)$, and the equilibrium will be stable if $\mathcal{E}(\overline{q}_1, ..., \overline{q}_d)$ is minimum.

Unfortunately, a stable equilibrium is impossible, due to a minimum property of the electrostatic energy first proven by Lord Kelvin.

Theorem 2.4.1 *(Kelvin's theorem.) The electric energy $\mathcal{E}_D$ is minimum in equilibrium (static) conditions, if the conductors' surface charges $Q_1, ..., Q_N$ are kept constant and if $\boldsymbol{E}$ satisfies the asymptotic condition*

$$\boldsymbol{E} = O(|\boldsymbol{x}|^{-2}) \quad \textit{as } |\boldsymbol{x}| \to \infty$$

Proof. The electric field $\boldsymbol{E}$ and displacement vector $\boldsymbol{D}$ satisfy the equations

$$\boldsymbol{D} = \epsilon \boldsymbol{E}\,, \qquad div\, \boldsymbol{D} = 0$$

in D. Moreover the Gauss Law says that

$$Q_j = \int_{S_j} \boldsymbol{D} \cdot \boldsymbol{n}\, dS \qquad\qquad (j = 1, ..., N)$$

where $S_1, ..., S_N$ are the conductors's surfaces with normal $\boldsymbol{n}$ oriented towards D. Let

$$\boldsymbol{E} = \boldsymbol{E}_o + \boldsymbol{E}' \quad , \qquad \mathrm{D} = \boldsymbol{D}_o + \boldsymbol{D}' = \epsilon \boldsymbol{E}_o + \epsilon \boldsymbol{E}'$$

where $\boldsymbol{E}_o$ is the electrostatic field corresponding to the given charges $Q_1, ..., Q_N$, i.e. the solution of Problem II in §2.3.1, and $\boldsymbol{E}'$ is not identically zero. Then

$$\boldsymbol{D} \cdot \boldsymbol{E} = \boldsymbol{D}_o \cdot \boldsymbol{E}_o + \boldsymbol{D}' \cdot \boldsymbol{E}' + 2\boldsymbol{D}' \cdot \boldsymbol{E}_o = \epsilon(|\boldsymbol{E}_o|^2 + |\boldsymbol{E}'|^2) + 2\boldsymbol{D}' \cdot \boldsymbol{E}_o$$

and the electric energy is

$$\begin{aligned}\mathcal{E} = \mathcal{E}_D &= \frac{1}{2}\int_D \boldsymbol{D} \cdot \boldsymbol{E}\, dV = \frac{\epsilon}{2}\int_D |\boldsymbol{E}_o|^2 dV + \int_D \boldsymbol{D}' \cdot \boldsymbol{E}_o\, dV + \frac{\epsilon}{2}\int_D |\boldsymbol{E}'|^2 dV \\ &> \frac{\epsilon}{2}\int_D |\boldsymbol{E}_o|^2 dV + \int_D \boldsymbol{D}' \cdot \boldsymbol{E}_o\, dV = \mathcal{E}_o + \int_D \boldsymbol{D}' \cdot \boldsymbol{E}_o\, dV\end{aligned}$$

where $\mathcal{E}_o$ is the electrostatic energy in D. By assumption we have

$$div\,\boldsymbol{D}_o = 0\ ,\ \ \boldsymbol{E}_o = -grad\,u$$

and, since in equilibrium (static) conditions the conductors' charges are the same,

$$Q_j = \int_{S_j} \boldsymbol{D}_o \cdot \boldsymbol{n}\,dS \quad , \qquad \int_{S_j} \boldsymbol{D}' \cdot \boldsymbol{n}\,dS \ = 0 \qquad\qquad (j = 1, ..., N)$$

The asymptotic condition at infinity implies then that

$$\int_D \boldsymbol{D}'\cdot\boldsymbol{E}_o\;dV = -\int_D \boldsymbol{D}'\cdot grad\,u\,dV = -\int_D div(\boldsymbol{D}'u)\,dV = -\int_{\partial D} u\boldsymbol{D}'\cdot\boldsymbol{n}\,dV$$

and u is constant on the conductors' surfaces $S_j\,(j = 1, ..., N)$, so that

$$\int_D \boldsymbol{D}' \cdot \boldsymbol{E}_o\;dV = 0$$

Thus $\mathcal{E}_D > \mathcal{E}_o$, as asserted.

Kelvin's theorem applies also to a fictitious electric field $\boldsymbol{E}$ (i.e. an electric field which cannot be realized in practice) provided all the equations for $\boldsymbol{D}$ and $\boldsymbol{E}$ stated in the proof are satisfied. Then if the charges are fixed, the fictitious electric energy is greater than the electrostatic energy.

Theorem 2.4.2 *(Earnshaw's Theorem.) A charged conductor cannot be held in stable equilibrium by the electrostatic forces arising from the presence of other charged conductors.*

Proof [27]. Consider an isolated system consisting of N conductors $C_1, ..., C_N$ with total surface charges $Q_1, ..., Q_N$ and suppose that the system is in equilibrium. We denote by $\sigma_1, ..., \sigma_N$ the surface charge densities on the conductors surfaces $S_1, ..., S_N$, respectively.

Let us take a cartesian coordinate system $O(x, y, z)$ with origin in the center of mass M_o of the free conductor C, which without loss of generality can be identified with C_1. Suppose that C undergoes a (rigid) translation so that the new position M of its center of mass has coordinates $\boldsymbol{a}= (a, b, c)$

with respect to $O = M_o$. This translation must be such that $C = C_1$ does not touch any other conductor $C_2, ..., C_N$. Let $M(x', y', z')$ be a new coordinate system with origin in M and axes parallel to the previous ones. We have then

$$x{=}a{+}x', \quad y{=}b{+}y', \quad z{=}c{+}z' \qquad \Leftrightarrow \qquad \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{x}' \tag{2.79}$$

Before proceeding, we remark that if $w(\boldsymbol{x})$ is a harmonic function in a domain D and if $\boldsymbol{x}'$ is a fixed point of D, then $w = w(\boldsymbol{a}{+}\boldsymbol{x}')$ is a harmonic function of $\boldsymbol{a}= (a, b, c)$ in the corresponding translated domain D'.

Let C' denote the translated conductor C_1 and S' its surface. After the translation the total surface charges $Q_1, ..., Q_N$ remain the same because of the conservation of charge, eq. (1.28), but in equilibrium (static) conditions the surface charge densities will in general be different, so as to guarantee that the conductors' surfaces are equipotential in the new configuration. Consider instead the fictitious electric field obtained by assuming that the surface charge densities $\sigma_1', ..., \sigma_N'$ after the translation are the same as before, i.e. are given by

$$\sigma_1' = \sigma_1(\boldsymbol{x}') \text{ on } S' \quad , \qquad \sigma_j' = \sigma_j(\boldsymbol{x}) \text{ on } S_j \quad (j = 2, ..., N)$$

where $\boldsymbol{x}$ is given by (2.79). The corresponding fictitious electric potential of the system can be written as the sum

$$u = u'(\boldsymbol{x}') + u''(\boldsymbol{x})$$

of the potential $u'(\boldsymbol{x}')$ due to the charge distribution σ_1' on the first conductor and the potential $u''(\boldsymbol{x})$ due to the charge distribution $(\sigma_2', ..., \sigma_N')$ on the remaining conductors. The potential u' depends only on $\boldsymbol{x}'$, since so does σ_1'.

For an arbitrarily fixed M consider the following function of $\boldsymbol{a}$:

$$\mathcal{E}'(\boldsymbol{a}) \;:=\; \frac{1}{2}\int_{S'} u'\sigma_1\, dS \;+\; \frac{1}{2}\int_{S'} u''\sigma_1\, dS \;+\; \frac{1}{2}\int_{S''} u'\sigma\, dS \;+\; \frac{1}{2}\int_{S''} u''\sigma\, dS$$

where $S'' = S_2 \cup S_3 ... \cup S_N$ and $\sigma = \sigma_j$ on S_j $(j = 2, ..., N)$. It is easy to see that:

(i) the first and the fourth integrals in the expression of $\mathcal{E}'(\mathbf{a})$ are independent of $\boldsymbol{a}$

(ii) the second and the third integrals depend on a through u' and u'', since

$$\int_{S'} u''(\boldsymbol{x})\ \sigma_1(\boldsymbol{y})\, dS_y = \int_{S'} u''(\boldsymbol{y}+\boldsymbol{a})\ \sigma_1(\boldsymbol{y})\, dS_y$$
$$\int_{S''} u'(\boldsymbol{x}')\, \sigma(\boldsymbol{y})\ dS_y = \int_{S''} u'(\boldsymbol{y}-\boldsymbol{a})\, \sigma(\boldsymbol{y})\ dS_y$$

(iii) if $\boldsymbol{a}= \mathbf{0}$, eq. (2.71) implies that $\mathcal{E}'(\mathbf{0})$ coincides with the electrostatic energy $\mathcal{E}_\mathrm{o}$ in the equilibrium configuration $M = M_o$.

Since u', u'' are harmonic functions, the above discussion shows that $\mathcal{E}'(\mathbf{a})$ is a harmonic function of $\boldsymbol{a}$ in a neighborhood of the origin, and hence cannot have a minimum at $\boldsymbol{a}= \mathbf{0}$ [2]. It follows that there exists a translation $\boldsymbol{a}$ from M_o to some point M such that

$$\mathcal{E}'(\mathbf{a}) < \mathcal{E}'(\mathbf{0})$$

and since $\mathcal{E}'(\boldsymbol{a})$ is the energy of a fictitious electric field, Kelvin's theorem implies that the electrostatic energy $\mathcal{E}$ in the configuration M, corresponding to a rearranged charge distribution on the conductors' surfaces, is less than $\mathcal{E}'$:

$$\mathcal{E} \ < \mathcal{E}' \ < \mathcal{E}_\mathrm{o}$$

Therefore $\mathcal{E} < \mathcal{E}_\mathrm{o}$ in the configuration M and the equilibrium is unstable, as asserted.

Earnshaw's theorem originally applied to systems of charged particles [9]. Thus for example a charge in the middle of a box in equilibrium with equal or opposite ones at the corners cannot be in stable equilibrium. This theorem is also of historical importance, as it implied the impossibility of constructing an electrical model of the atom with the nucleus and the electrons represented by charged particles at rest under the mutual electric interactions.

2.5 Magnetic field of a permanent magnet

Consider a bar magnet $\mathcal{M}$ in the shape of a right circular cylinder with directrices parallel to the $z = x_3$ axis, height $h = 2l$ and circular bases Σ : $z = -l$ and Σ' : $z = l$ with radius $R \ll l$ (see Fig. 2.3). Suppose $\mathcal{M}$ is

[9] W. Earnshaw, Trans. Camb. Phil. Soc. 7, 97-112, 1842

surrounded by empty space, or by any homogeneous dielectric with magnetic permeability $\mu = \mu_o$. We denote here $\boldsymbol{x}= (x, y, z)$.

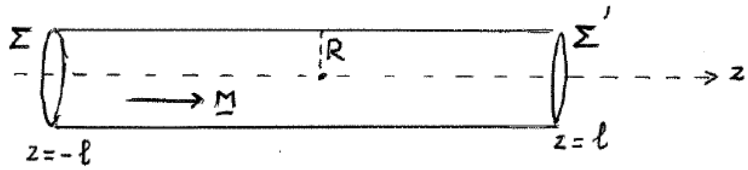

Figure 2.3: Bar magnet

The amount of magnetization inside $\mathcal{M}$ is measured by the magnetization vector

$$\boldsymbol{M} := \frac{1}{\mu_o}\boldsymbol{B} - \boldsymbol{H} \tag{2.80}$$

(eq. (1.83)), whereas outside the magnet $\boldsymbol{B}= \mu_o \boldsymbol{H}$ and the magnetization vector is identically zero. The experimental evidence shows that for $R \ll l$, $\boldsymbol{M}$ can be approximated by a constant vector parallel to the $z-$axis in $\mathcal{M}$,

$$\boldsymbol{M} \cong M_o I_{\mathcal{M}}(\boldsymbol{x})\mathbf{c}_3 \quad , \quad I_{\mathcal{M}}(\boldsymbol{x}) := \begin{cases} 1 & \boldsymbol{x} \in \overline{\mathcal{M}} \\ 0 & \text{otherwise} \end{cases}$$

As a first step of approximation, we begin by taking the solenoidal vector $\boldsymbol{B}$ constant and parallel to the $z-$axis inside $\mathcal{M}$

$$\boldsymbol{B} \cong B\mathbf{c}_3 \qquad \text{for } \sqrt{x^2 + y^2} \le R, \ -l \le z \le l$$

The previous equations imply that a first approximation for the magnetic field $\boldsymbol{H}= \mu^{-1}\boldsymbol{B}-\boldsymbol{M}$ is

$$\boldsymbol{H} \cong H\mathbf{c}_3 = \mu_o^{-1}\boldsymbol{B}(\boldsymbol{x}) - M_o I_{\mathcal{M}}(\boldsymbol{x})\mathrm{c}_3 \equiv \begin{cases} (\mu_o^{-1}B - M_o)\mathbf{c}_3 & in\mathcal{M} \\ \mu_o^{-1}\boldsymbol{B}(\boldsymbol{x}) & outside\mathcal{M} \end{cases}$$

Denoting by $-$ the interior of $\mathcal{M}$ and by $+$ the exterior, the matching relations (R2) and (R4) in §1.4, written for $\mathbb{S} = \partial\mathcal{M}$, yield

$$[\boldsymbol{B}]_{\partial\mathcal{M}} \cdot \boldsymbol{n} = 0 \quad \Rightarrow \quad [\boldsymbol{H}]_{\partial\mathcal{M}} \mathrm{c}_3 \cdot \boldsymbol{n} = -M_o [I_{\mathcal{M}}]_{\partial\mathcal{M}} \mathbf{c}_3 \cdot \boldsymbol{n}$$
$$[\boldsymbol{H}]_{\partial\mathcal{M}} \wedge \boldsymbol{n} = \boldsymbol{0} \quad \Rightarrow \quad [\boldsymbol{B}]_{\partial\mathcal{M}} \mathbf{c}_3 \wedge \boldsymbol{n} = \mu_o M_o [I_{\mathcal{M}}]_{\partial\mathcal{M}} \mathbf{c}_3 \wedge \boldsymbol{n}$$

where the magnet boundary $\partial\mathcal{M}$ consists of the two bases and of the lateral surface S_L.

Consider first the two bases Σ and Σ', whose normals are $\boldsymbol{n} = -\mathbf{c}_3$ and $\boldsymbol{n} = \mathbf{c}_3$, respectively. The second condition is automatically satisfied, since $\mathbf{c}_3 \wedge \boldsymbol{n} = \mathbf{0}$; the first shows that $\boldsymbol{B}$ is continuous across the two bases and, with the usual notations, we have

$$[\boldsymbol{B}]_{\Sigma} = 0 \quad \Rightarrow \quad H_{-} - H_{+} = -M_o \qquad \text{across } \Sigma : z = -l \ , \ \boldsymbol{n} = -\mathbf{c}_3$$
$$[\boldsymbol{B}]_{\Sigma'} = 0 \quad \Rightarrow \quad H'_{+} - H'_{-} = M_o \qquad \text{across } \Sigma' : z = l \ , \ \boldsymbol{n} = \mathbf{c}_3$$

Since at this level of approximation $H'_- = H_- = H$, the previous relations yield

$$H'_{+} = H_{+} = \frac{M_o}{2} \quad ; \quad H'_{-} = H_{-} = H = -\frac{M_o}{2}\,, \quad B = \mu_o \frac{M_o}{2} \qquad \text{in } \mathcal{M} \tag{2.81}$$

and as $[\boldsymbol{H}\cdot\boldsymbol{n}]_{\Sigma} = H_- - H_+ = -[\boldsymbol{H}\cdot\boldsymbol{n}]_{\Sigma'}$ we have

$$[\boldsymbol{H}] \cdot \boldsymbol{n} = \begin{cases} M_o & \text{across } \Sigma' \ : \ z = l \\ -M_o & \text{across } \Sigma \ : \ z = -l \end{cases} \tag{2.82}$$

Across the lateral surface S_L, where $\mathbf{c}_3 \cdot \boldsymbol{n} = 0$, the first matching relation is identically satisfied, the second shows that $\boldsymbol{H}$ is continuous across S_L, whereas $\boldsymbol{B}$ is discontinuous:

$$[\boldsymbol{H}]_{S_L} = \mathbf{0} \quad \Rightarrow \quad B_{+} - B_{-} = -\,\mu_o M_o \ \ \text{across } S_L \tag{2.83}$$

Eq. (2.82) can now be used in order to improve the approximations of $\boldsymbol{H}$ and $\boldsymbol{B}$: namely, since $\boldsymbol{B}$ is solenoidal everywhere and $\boldsymbol{M}$ is solenoidal in side the magnet, we can assume as a second approximation step that $\boldsymbol{H}$ is an irrotational and solenoidal vector field in all $\mathbb{R}^3$ except the two bases, where it satisfies the jump relations (2.82). In other words, we assume that

$$\boldsymbol{H} = -grad\, v$$

where the magnetic potential v satisfies the boundary value problem

$$\begin{aligned} &\Delta_3 v = 0 && \text{in } \mathbb{R}^3 \setminus \Sigma \cup \Sigma' \\ &[\frac{\partial v}{\partial n}]_{\Sigma} = M_o\,, \quad [\frac{\partial v}{\partial n}]_{\Sigma'} = -M_o && \\ &v = O(|\boldsymbol{x}|^{-1}) && \text{as } |\mathbf{x}| \to +\infty \end{aligned} \tag{2.84}$$

The analysis carried out in §2.1, and in particular eq. (2.19), shows that the solution is given by the superposition of the two single layer potentials over Σ and Σ' with constant densities $\pm M_o$:

$$v(\boldsymbol{x}) = \frac{M_o}{4\pi}\int_{\Sigma'}\frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|} - \frac{M_o}{4\pi}\int_{\Sigma}\frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|} \tag{2.85}$$

Eq. (2.80) yields then the corresponding upgraded approximation for the magnetic induction

$$\boldsymbol{B}(\boldsymbol{x}) = \mu_o M_o I_{\mathcal{M}}(\boldsymbol{x})\mathbf{c}_3 - grad\, v(\boldsymbol{x})$$

If $\boldsymbol{y}_1, \boldsymbol{y}_2$ are two suitable points belonging to Σ and Σ', respectively, the mean value theorem applied to the integrals in eq. (2.85) says that

$$v(\boldsymbol{x}) = \frac{M_o R^2}{4|\boldsymbol{x}-\boldsymbol{y}_2|} - \frac{M_o R^2}{4|\boldsymbol{x}-\boldsymbol{y}_1|}$$

and far from $\mathcal{M}$, for $|\boldsymbol{x}| \gg l$, v approaches the potential (2.11)

$$v(\boldsymbol{x}) = \frac{\boldsymbol{m}\cdot\boldsymbol{x}}{4\pi|\boldsymbol{x}|^3}$$

due to a magnetic dipole with moment $\boldsymbol{m}= 4\pi l M_o R^2\mathbf{c}_3$ placed at the origin (Exercise 23).

On the other hand, eq. (2.85) written for $\boldsymbol{x} = \boldsymbol{y} = 0$ yields the value of the magnetic field on the cylinder axis

$$\boldsymbol{H}(0,0,z) = H(z)\mathbf{c}_3$$

with

$$H(z) := -\frac{\partial v(0,0,z)}{\partial z} = \frac{M_o}{2}\left\{\frac{l-z}{\sqrt{(l-z)^2+R^2}} + \frac{l+z}{\sqrt{(l+z)^2+R^2}} - 2\right\} \tag{2.86}$$

(Exercise 24). The limiting values as $z \to \pm l$ must coincide by construction with those obtained at the previous step, eq. (2.81). Indeed, it is easy to see that

$$H'_- := \lim_{z\to l-0} H(z) = -\frac{M_o}{2} \qquad H'_+ := \lim_{z\to l+0} H(z) = \frac{M_o}{2}$$

and since $H(z)$ is even in z, we have

$$H_- = H'_- = -\frac{M_o}{2} \quad , \quad H'_+ = H_+ = \frac{M_o}{2} \tag{2.87}$$

as before. However, at the present level of approximation, $\mathbf{H}$ (and hence $\boldsymbol{B}$) are no longer constant inside the magnet. In particular, near the center of $\mathcal{M}$ we can neglect the R^2-terms in the denominators of eq. (2.86) and for $|z| \ll l$ we obtain

(2.88)

$$\begin{aligned} \boldsymbol{H}(0,0,z) &\sim \frac{M_o}{2}\left\{\frac{l-z}{\sqrt{l^2-2lz}} + \frac{l+z}{\sqrt{l^2+2lz}} - 2\right\}\mathbf{c}_3 \sim -M_o\frac{z^2}{l^2}\,\mathbf{c}_3 \cong 0 \\ \boldsymbol{B}(0,0,z) &\sim -\mu_o M_o\frac{z^2}{l^2}\,\mathbf{c}_3 + \mu_o M_o\mathbf{c}_3 \sim \mu_o M_o\mathbf{c}_3 \end{aligned}$$

whereas near the bases inside the magnet, i.e. for $z \cong \pm l$ we have

$$\begin{aligned} \boldsymbol{H}(0,0,z) &\sim H_-\mathbf{c}_3 = -\frac{M_o}{2}\mathrm{c}_3 \\ \boldsymbol{B}(0,0,z) &\sim -\mu_o\frac{M_o}{2}\mathrm{c}_3 + \mu_o M_o\mathrm{c}_3 \sim \mu_o\frac{M_o}{2}\mathrm{c}_3 \end{aligned} \tag{2.89}$$

Since $R \ll l$ by assumption, the values of $\boldsymbol{H}$ and $\boldsymbol{B}$ in the magnet are close to those on the magnet axis $\boldsymbol{x} = \boldsymbol{y} = \mathbf{0}$. We conclude that, at this level of approximation, we have inside the magnet

$$\boldsymbol{H} = H(z)\mathbf{c}_3 \quad , \boldsymbol{B} = B(z)\mathbf{c}_3 \ , \quad \boldsymbol{M} = M_o\mathbf{c}_3 \tag{2.90}$$

with $H(z) \cong 0$ for $|z| \ll l$, and

$$H(z) < 0 \quad , \quad B(z) > 0 \qquad\qquad \text{for } |z| \leq l$$

(see Exercise 24). Thus the direction of $\boldsymbol{H}$ opposes that of $\boldsymbol{B}$ inside the magnet, so that the field $\boldsymbol{H}$ is demagnetizing. This demagnetizing character of the magnetic field in a permanent magnet might be inferred from the Maxwell equations and the fact that all the lines of force of $\boldsymbol{H}$ (with the exception of the $z-$axis) are closed. If Γ is such a closed line of force, the Ampère circuital law (1.25) of Chapter 1 with $\boldsymbol{J}$=$\boldsymbol{D}$= 0

$$\oint_\Gamma \boldsymbol{H}\cdot\boldsymbol{t}\,ds = 0$$

implies that $\boldsymbol{H}$ must be discontinuous along Γ and must oppose $\boldsymbol{B}$ (and $\boldsymbol{M}$) inside the magnet. The permanent magnetization thus corresponds to that portion of the hysteresis loop where $\boldsymbol{B}$ and $\boldsymbol{H}$ like $\boldsymbol{M}_o$ and $\boldsymbol{H}$, have opposite signs (see Chapter 1, §1.8).

Exercises

Exercise 1. Consider the summation problem in a homogeneous dielectric $D = \mathbb{R}^3$ with permittivity ϵ

$$\text{(E1)} \qquad \Delta_3 u(\boldsymbol{x}) = 0 \qquad \text{for } r \neq R \quad ; \quad \epsilon \left[(\frac{\partial u}{\partial n})_+ - (\frac{\partial u}{\partial n})_- \right] = -\sigma_o$$

where $r = |\boldsymbol{x}|$, $\sigma_o \neq 0$ is a constant surface charge over the sphere $r = R$ with outer normal $\boldsymbol{n}$. Suppose $u(\boldsymbol{x})$ is continuous in $\mathbb{R}^3$ and satisfies condition (2.5) at infinity: $u(\boldsymbol{x})$ depends only on r for reasons of symmetry and is given by the radial harmonic functions $u_+ := A/|\boldsymbol{x}|$ for $r > R$ and $u_- := A/R$ for $r \leq R$. We have

$$(\frac{\partial u}{\partial n})_\pm \;=\; \lim_{r \to R \pm 0} \frac{du}{dr}$$

and since $(\frac{\partial u}{\partial n})_- = 0$, $(\frac{\partial u}{\partial n})_+ = -A/R^2$, the second equation (E1) coincides with (2.3) if $A = \sigma_o R^2/\epsilon$. We conclude that the potential

$$u(\boldsymbol{x}) \;=\; \begin{cases} \frac{\sigma_o R}{\epsilon} = u_- & r \leq R \\ \frac{\sigma_o R^2}{\epsilon |\boldsymbol{x}|} = u_+ & r > R \end{cases}$$

solves the summation problem (E1). Since $u(\boldsymbol{x})$ is constant for $r \leq R$, this is also the solution for a uniformly charged conducting ball $r < R$ having the potential $\frac{\sigma_o R}{\epsilon}$ surrounded by the unbounded dielectric $D : r > R$. Since $\Delta_3 \, u_\pm = 0$ and

$$0 = \int_{r<R} \Delta_3 \, u_- \, dV \;=\; - \int_{\partial D} (\frac{\partial u}{\partial n})_- \, dS$$

$$0 = \int_{r>R} \Delta_3 \, u_+ \, dV \;=\; - \int_{\partial D} (\frac{\partial u}{\partial n})_+ \, dS - \frac{\sigma_o}{\epsilon} = 0$$

eq. (2.6) is true for the bounded ball $r < R$ and for the unbounded domain $r > R$. The electric field $\boldsymbol{E} = -grad\, u$ is given by

$$\boldsymbol{E}(\boldsymbol{x}) = \begin{cases} \boldsymbol{0} & for\,|\boldsymbol{x}| < R \\ \frac{Q_{tot}\,\boldsymbol{x}}{4\pi\epsilon|\boldsymbol{x}|^3} & for|\boldsymbol{x}| > R \end{cases}$$

where $Q_{tot} = 4\pi\sigma_o R^2$ is the total charge contained in any sphere with radius larger than R, in accordance with the Gauss Law (1.4)

Note that one may take $u_- := B$ for $r \leq R$, where B is any constant. Then u is discontinuous and this discontinuity may be thought of as being due to the presence of a double layer potential of constant density distributed over the sphere (see §2.1.2).

Exercise 2. Consider the summation problem for a homogeneous dielectric $D = \mathbb{R}^3$ with a piecewise constant volume charge

$$\rho = \begin{cases} \rho_o & \text{for } |\boldsymbol{x}| < R \\ 0 & \text{for } |\boldsymbol{x}| > R \end{cases}$$

Suppose that u is continuous and satisfies (2.5) at infinity. Since $u = u(r)$ we have (see previous exercise)

$$u_+(r) = \frac{A}{r} \qquad \text{for } r = |\boldsymbol{x}| > R$$

(radial harmonic function), while from eq. (2.1) we have

$$\frac{1}{r}\frac{d^2}{dr^2}(ru(r)) = -\frac{\rho_o}{\epsilon} \qquad \Rightarrow \qquad u_-(r) = B - \frac{\rho_o}{6\epsilon}r^2 \qquad \text{for } \; r < R$$

The constant B can be computed in terms of A and ρ_o from the continuous matching condition $u_+ = u_-$ for $r = R$:

$$B = \frac{A}{R} + \frac{\rho_o}{6\epsilon}R^2$$

The surface charge density $\sigma = \sigma_o$ is constant on the sphere $r = R$, with

$$\sigma_o := -\epsilon\frac{du_+}{dr} + \epsilon\frac{du_-}{dr}\Big|_{r=R} = \epsilon\frac{A}{R^2} - \epsilon\frac{\rho_o}{3\epsilon}R$$

so that σ_o is determined by the trace $u_+ = A/R$ over the sphere and by ρ_o . Conversely, if we assign σ_o and ρ_o we can determine the constant A in the form

$$A = \frac{\sigma_o}{\epsilon} R^2 + \frac{\rho_o}{3\epsilon} R^3 \equiv \frac{Q_{tot}}{4\pi\epsilon}$$

where $Q_{tot} = 4\pi\sigma_o R^2 + \frac{4}{3}\pi R^3 \rho_o$ is the total charge contained in any sphere with radius larger than R. The electric field $\boldsymbol{E} = -grad\, u$ is then given by

$$\boldsymbol{E}(\boldsymbol{x}) = \begin{cases} \frac{\rho_o}{3\epsilon}\boldsymbol{x} & for\ |\boldsymbol{x}| < R \\ \frac{Q_{tot}\,\boldsymbol{x}}{4\pi\epsilon|\boldsymbol{x}|^3} & for\ |\boldsymbol{x}| > R \end{cases}$$

in accordance with the Gauss Law (1.4)

Exercise 3. For reasons of symmetry the electric field $\boldsymbol{E}$ depends only on the normal coordinate z and its lines of force are straight lines orthogonal to the plane, so that $\boldsymbol{E} = E(z)\boldsymbol{n} = -du/dz$. Since $\rho \equiv 0$ and ϵ is constant we have

$$0 = div(\boldsymbol{D}) = \epsilon div\, \boldsymbol{E} = \epsilon \frac{dE}{dz}$$

and so $\boldsymbol{E} = E_o \mathbf{c}_3 = E_o \boldsymbol{n}$ is a constant vector. The Gauss law applied to a cylindrical pillbox $\Omega_h = \mathbb{I}_o \times (-h, h)$ with generatrices parallel to $\boldsymbol{n}$ and section area A_o (§1.4) yields $\epsilon E_o A_o = \sigma A_o$, whence eqs. (2.11) and (2.12) follow.

An alternative approach is to solve the Neumann problem for the half-space

$$\frac{d^2u}{dz^2} = 0 \quad \text{for } z > 0 \quad , \quad \frac{du}{dz} = -\frac{\sigma}{\epsilon} \quad \text{for } z = 0$$

(cfr. eq. (2.3)), whose solution is $u = -\ \sigma z/\epsilon + ç$.

Exercise 4. Solve the analagous problem for the case when the half-space $z < 0$ is a dielectric with the same dielectric constant ϵ as $z > 0$.

$$\text{Answer:} \quad u = \begin{cases} -\sigma z/2\epsilon & z > 0 \\ \sigma z/2\epsilon & z < 0 \end{cases} \quad , \quad \boldsymbol{E} = \begin{cases} \sigma \boldsymbol{n}/2\epsilon & z > 0 \\ -\sigma \boldsymbol{n}/2\epsilon & z < 0 \end{cases}$$

Exercise 5. Since

$$|\boldsymbol{x} - \boldsymbol{y}| = \sqrt{(\boldsymbol{x} - \boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})} = \sqrt{|\boldsymbol{x}|^2 - 2\boldsymbol{y} \cdot \boldsymbol{x} + |\boldsymbol{y}|^2}$$

letting $|\boldsymbol{x}| \to +\infty$ we have

$$\frac{1}{|\boldsymbol{x} - \boldsymbol{y}|} = \frac{1}{|\boldsymbol{x}|\sqrt{1 - (2\boldsymbol{y} \cdot \boldsymbol{x} + |\boldsymbol{y}|^2)/|\boldsymbol{x}|^2}} = \frac{1}{|\boldsymbol{x}|}(1 + \frac{2\boldsymbol{y} \cdot \boldsymbol{x}}{|\boldsymbol{x}|^2} + O(\frac{1}{|\boldsymbol{x}|^2})) = \frac{1}{|\boldsymbol{x}|} + O(\frac{1}{|\boldsymbol{x}|^2})$$

Substituting in (2.14) we get eq. (2.15). Similarly for the single layer potential.

Exercise 6. For any point $\boldsymbol{x}_o \in \partial\mathbb{S}$, let $\boldsymbol{t}(\boldsymbol{x}_o)$ be the unit tangent vector to $\partial\mathbb{S}$ and $\boldsymbol{n}(\boldsymbol{x}_o)$ the unit normal to $\mathbb{S}$ at the point $\boldsymbol{x}_o$. Verify that the gradient of the single layer potential satisfies

$$grad_x \int_{\mathbb{S}} \frac{\sigma(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\, dS_y \;\sim\; 2\sigma(\boldsymbol{x}_o)\boldsymbol{t}(\boldsymbol{x}_o) \wedge \boldsymbol{n}(\boldsymbol{x}_o)\, log|\boldsymbol{x}-\boldsymbol{x}_o| \;+\; O(1)$$

as $\boldsymbol{x} \rightarrow \boldsymbol{x}_o$, where $\boldsymbol{t} \wedge \boldsymbol{n}$ is tangent to $\mathbb{S}$ and orthogonal to $\partial\mathbb{S}$.

Exercise 7. If $\mathbb{S}$ is the rectangle $(0, A) \times (0, B)$ in the $(y_2, y_3)-$plane, $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{y} = (y_1, y_2, y_3)$, and $\sigma = 4\pi\epsilon$, the single layer potential $\mathcal{V} = \mathcal{V}_{4\pi}(\boldsymbol{x})$ is given by

$$\mathcal{V}(x_1, x_2, x_3) := \int_{\mathbb{S}} \frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|} \;=\; \int_0^A log\Big[\frac{B - x_3 + \sqrt{Z^2 + (x_2 - y_2)^2}}{-x_3 + \sqrt{V^2 + (x_2 - y_2)^2}}\Big] dy_2$$

where

$$Z := \sqrt{x_1^2 + (B - x_3)^2}\ , \qquad V := \sqrt{x_1^2 + x_3^2}$$

It follows that

$$\frac{\partial \mathcal{V}}{\partial x_3} = log\Big[\frac{-x_2 + \sqrt{Z^2 + x_2^2}}{-x_2 + \sqrt{V^2 + x_2^2}}\ \frac{A - x_2 + \sqrt{V^2 + (A - x_2)^2}}{A - x_2 + \sqrt{Z^2 + (A - x_2)^2}}\Big]$$

has logarithmic singularities on the horizontal sides of the rectangle $x_1 = x_3 = 0,\ x_1 = x_3 - B = 0$ for $0 < x_2 < A$,and is regular on the vertical sides $x_1 = x_2 = 0,\ x_1 = x_2 - A = 0$ for $0 < x_3 < B$, in accordance with Exercise 6.

Exercise 8. If $\boldsymbol{x}, \boldsymbol{y}$ vary over a plane surface $\mathbb{S}$ the vector $\boldsymbol{x}-\boldsymbol{y}$ is orthogonal to $\boldsymbol{n}(\boldsymbol{y})$ and so the integrand in $\mathcal{W}_{\nu_o}(\boldsymbol{x})$

$$\frac{\partial}{\partial n_y}\Big(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\Big) \equiv \frac{\boldsymbol{n}(\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3}$$

is zero for all $\boldsymbol{x} \neq \boldsymbol{y}$.

Exercise 9. Hint: The solid angle with sign $\varpi(\mathbf{x})$ subtended at $\mathbf{x}$ is 4π (total solid angle) if $\mathbf{x}$ is an interior point of Ω, 2π if $\mathbf{x}$ is on the boundary $\partial\Omega$, 0 (by cancellation) if $\mathbf{x}$ is outside $\overline{\Omega}$.

Exercise 10. Prove by direct calculation that the double layer potential $W_\pi(\boldsymbol{x})$ of density π, distributed over the coordinate half-plane $y_1 < 0,\ y_2 = 0,\ y_3 \in \mathbb{R}$, oriented with normal $\boldsymbol{n}$=$\mathbf{c}_2$ along the positive y_2-axis, coincides with the angle

$$\varphi := arctan(\frac{x_2}{x_1}) \qquad\qquad (-\pi \le \varphi \le \pi)$$

between the vector $\mathbf{X}= (x_1,\ x_2)$ and the y_1-axis. If $\boldsymbol{\tau}$ denotes the transverse unit vector in the $(x_1, x_2)-$plane, the gradient of $W_\pi(\boldsymbol{x})$ is given by the Biot-Savart expression

$$grad\, W_\pi(\boldsymbol{x}) = \frac{1}{|\mathbf{X}|}\ \boldsymbol{\tau}$$

and is singular at $\mathbf{X}= \mathbf{0}$. *Hint* : We have

$$W_\pi(\boldsymbol{x}) = \frac{1}{4}\int_{-\infty}^{0} dy_1 \int_{-\infty}^{+\infty} \mathrm{c}_2 \cdot grad_y \frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\ dy_3$$

where $|\boldsymbol{x}-\boldsymbol{y}| = \sqrt{(x_1-y_1)^2 + x_2^2 + (x_3-y_3)^2}$, whence

$$\mathbf{c}_2 \cdot grad_y \frac{1}{|\boldsymbol{x}-\boldsymbol{y}|} = \frac{x_2}{|\boldsymbol{x}-\boldsymbol{y}|^3} = \frac{x_2}{\Big[(x_1-y_1)^2 + x_2^2 + (x_3-y_3)^2\Big]^{3/2}}$$

The integral on y_3 can be calculated by means of the change of variables

$$\eta = arcsinh\,(\frac{y_3 - x_3}{q}) \quad , \quad q = \sqrt{(x_1-y_1)^2 + x_2}$$

and the remaining integral on y_1 by means of the change of variables $t = exp\,(arcsinh(P))\,, \quad P = \frac{y_1 - x_1}{x_2}$.

Exercise 11. Consider the sum of two double layer potentials

$$u(\boldsymbol{x}) := \frac{\nu_i}{4\pi}\int_{S_i} \frac{\partial}{\partial n_y}(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|})\ dS_y\ + \frac{\nu_e}{4\pi}\int_{S_e} \frac{\partial}{\partial n_y}(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|})\ dS_y$$

where the closed bounded surfaces $S_i,\ S_e$ are the interior and exterior boundary, respectively, of a normal domain Ω, $S_i = \partial\Omega_i$, with the normal $\boldsymbol{n}$ to S_i

oriented towards the interior of Ω_i and the normal $\boldsymbol{n}$ to S_e towards the exterior of Ω. Then

$$u(\boldsymbol{x}) = \begin{cases} \nu_i - \nu_e & \boldsymbol{x} \in \Omega_i \\ \frac{1}{2}\nu_i - \nu_e & \boldsymbol{x} \in S_i \\ -\nu_e & \boldsymbol{x} \in \Omega \\ -\frac{1}{2}\nu_e & \boldsymbol{x} \in S_e \\ 0 & \boldsymbol{x} \in \mathbb{R}^3 \backslash \overline{\Omega_i \cup \Omega} \end{cases}$$

Thus $\boldsymbol{E} = -grad\, u(\boldsymbol{x})$ is (almost) everywhere zero, the normal derivative of u vanishes over S_i and S_e, and if $\nu_i = \nu_e = \nu_o$ we obtain formally the same results as in eq. (2.27).

Exercise 12 (Green's first and second identity [2]). (i) Let f, g be two biregular functions (i.e. $f, g \in C^2(\Omega) \cap C^1(\overline{\Omega})$) in a normal domain $\Omega \subset \mathbb{R}^3$ with outer normal $\boldsymbol{n}$ to $\partial\Omega$. Green's first identity for Ω says that

$$\int_\Omega f \Delta_3 g \; dV \; = \int_{\partial\Omega} f \frac{\partial g}{\partial n} \, dS - \int_\Omega grad\, f \cdot grad\, g \; dV$$

In particular if $f = g$

$$\int_\Omega f \Delta_3 f \, dV \; = \int_{\partial\Omega} f \frac{\partial f}{\partial n} \; dS - \int_\Omega |grad\, f|^2 \, dV$$

and if f is harmonic in Ω

(E2) $$\int_{\partial\Omega} f \frac{\partial f}{\partial n} \; dS = \int_\Omega |grad\, f|^2 \, dV$$

(ii) Let f, g be two biregular functions regular at infinity in the external domain $\Omega' = \mathbb{R}^3 \backslash \overline{\Omega}$, with $\mathbf{n}$ oriented as before on $\partial\Omega' = \partial\Omega$. Green's first identity for Ω' says that

$$\int_{\Omega'} f \Delta_3 g \, dV \; = -\int_{\partial\Omega'} f \frac{\partial g}{\partial n} \; dS - \int_{\Omega'} grad\, f \cdot grad\, g \, dV$$

In particular if $f = g$

$$\int_{\Omega'} f \Delta_3 f \, dV \; = -\int_{\partial\Omega'} f \frac{\partial f}{\partial n} \; dS - \int_{\Omega'} |grad\, f|^2 \, dV$$

and if f is harmonic in Ω'

$$\text{(E3)} \qquad -\int_{\partial\Omega'} f\frac{\partial f}{\partial n}\, dS = \int_{\Omega'} |grad\, f|^2\, dV$$

(*iii*) Interchanging $f \leftrightarrow g$ in Green's first identity for both Ω and Ω' yields Green's second identity for harmonic functions

$$\text{(E4)} \qquad \int_{\partial\Omega} (f\frac{\partial g}{\partial n} - g\frac{\partial f}{\partial n})\, dS \,=\, 0$$

Hint: use the identity

$$\text{(E5)} \qquad div\,(f grad\, g) \equiv f\,\Delta_3\, g \,+ grad\, f\cdot grad\, g$$

and the divergence theorem applied to Ω or, in case (*ii*), to the normal subdomain Ω_R of Ω' contained in a large sphere Σ_R , letting $R \to \infty$.

Exercise 13. Take $\chi = \chi_o$ constant, $\Gamma = \{-1 < y_2 < 1,\, y_1 = y_3 = 0\}$ *and* $\mathbf{x}=(0,0,z)$. *Then* $r=|\mathbf{x}-\mathbf{y}|=\sqrt{z^2+y_2^2}$ and for $z \neq 0$ the line potential

$$u(0,0,z) = \frac{\chi_o}{2\pi\epsilon}\int_0^1 \frac{dy}{\sqrt{z^2+y^2}} = \frac{\chi_o}{2\pi\epsilon}\int_0^{1/|z|} \frac{dt}{\sqrt{1+t^2}}$$

$$= \frac{\chi_o}{2\pi\epsilon} log\, \frac{\sqrt{z^2+1}-1}{|z|} \underset{z\to 0}{\longrightarrow} \frac{\chi_o}{2\pi\epsilon}(log|z|-log 2)$$

diverges logarithmically as $z \to 0$, that is for $\mathbf{x}\to\Gamma$. Hint: apply the substitution $y_2 = |z|t$ and the identity

$$arcsinh(\,\alpha\,) \,\equiv log\left[-\alpha + \sqrt{1+\alpha^2}\right]$$

Exercise 14. Since u_V and Vu_1 are both solutions of (2.41) regular at infinity, the uniqueness theorem for the exterior Dirichlet problem implies that $u_V = Vu_1$.

Exercise 15. The capacitary potential u_1 for a sphere of radius R centered at the origin must be a radial harmonic function taking the values zero at $r=\infty$ and 1 at $r = R$; hence

$$u_1 = R/r \quad , \quad r = |\mathbf{x}|$$

The corresponding surface charge density, given by eq. (2.44),

$$\sigma_1 = -\epsilon \frac{\partial u_1}{\partial n} = \epsilon R \frac{\partial}{\partial r}\frac{1}{r}\Big|_{r=R} = \frac{\epsilon}{R}$$

is constant on the sphere, so that the capacity is

$$\mathbb{C} = Q = 4\pi R^2 \frac{\epsilon}{R} = 4\pi\epsilon R \tag{E6}$$

Exercise 16 (homogeneous sphere in a uniform electric field). Consider a sphere of radius R surrounded by a dielectric medium of constant permittivity $\epsilon_+ = \epsilon$ in the presence of a uniform electric field $\boldsymbol{E}_o = E_o \mathbf{c}_3$.

(i) For a conducting sphere the electrostatic potential in spherical coordinates (r, θ, φ) is given by

$$u = V\frac{R}{r} - E_o x_3 + E_o R^3 \frac{x_3}{r^3} \equiv V\frac{R}{r} - E_o r cos\theta + E_o R^3 \frac{cos\theta}{r^2} \qquad (r \geq R)$$

(independent of φ), and $u = V$, $\boldsymbol{E} = \mathbf{0}$ inside the sphere. The surface charge is

$$\sigma \equiv -\epsilon \frac{\partial u}{\partial r} = \epsilon \frac{V}{R} + 3E_o\, cos\theta \quad \Rightarrow \quad Q = 4\pi\epsilon R V \equiv \mathbb{C}V$$

(see Exercise 15). The value of V must be specified in advance (Problem A).

(ii) For a dielectric sphere with permittivity ϵ_- we have

$$u = \begin{cases} u_- := -bE_o r cos\theta & (r < R) \\ u_+ := -E_o r cos\theta + E_o a\, cos\theta/r^2 & (r \geq R) \end{cases}$$

$$a = R^3 \frac{\epsilon_- - \epsilon_+}{\epsilon_- + 2\epsilon_+}\ , \ b = \frac{3\epsilon_+}{\epsilon_- + 2\epsilon_+}$$

and the surface charge is given by eq. (2.3) as

$$\sigma = -\epsilon_+ (\frac{\partial u}{\partial n})_+ + \epsilon_- (\frac{\partial u}{\partial n})_- = 2E_o\, \epsilon_+ \frac{\epsilon_- - \epsilon_+}{\epsilon_- + 2\epsilon_+} cos\theta$$

Here $R/r = u_1$ is the capacitary potential of the sphere (Exercise 15) and, by force of eq. (2.11), $cos\theta/r^2$ is the potential of a dipole placed in the center of the sphere with moment $\boldsymbol{m}$ parallel to $\mathbf{E}_o$. The lines of force of $\boldsymbol{E}$ are orthogonal to the sphere in case (i).

Exercise 17. Let D_ε denote the truncated (normal) domain bounded by ∂D and by a sphere S_ε with center $\mathbf{y}$, radius $\varepsilon > 0$ small enough and outer normal $\boldsymbol{n}$. Since the functions $G(\boldsymbol{x}|\boldsymbol{y})$ and $(4\pi|\boldsymbol{x}-\boldsymbol{y}|)^{-1}$ are harmonic in D_ε and

$$-\frac{\partial}{\partial n_x}(\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}) = \frac{\boldsymbol{n}(\boldsymbol{x})\cdot(\boldsymbol{x}-\boldsymbol{y})}{4\pi r^3} \equiv \frac{1}{4\pi r^2} \qquad (\forall \boldsymbol{y} \in D)$$

eq. (2.6) implies that

$$\int_{\partial D} \frac{\partial}{\partial n_x}(\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|})\, dS_x = -\int_{S_\varepsilon} \frac{\partial}{\partial n_x}(\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|})\, dS_x = 1 \qquad (\forall \boldsymbol{y} \in D)$$

But $g(\boldsymbol{x}|\boldsymbol{y})$ is harmonic in all of D, so that, using eq. (2.6) again we find

$$\int_{S_\varepsilon} \frac{\partial g(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} dS_x = \int_{\partial D} \frac{\partial g(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x}\, dS_x = 0 \qquad (\forall \boldsymbol{y} \in D)$$

and

$$\int_{\partial D} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x}\, dS_x = -\int_{S_\varepsilon} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} dS_x = -\int_{S_\varepsilon} \frac{\partial}{\partial n_x}(\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|})\, dS_x = 1$$

Finally $\frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} \geq 0$ follows from Proposition 2.2.6 (ii).

Exercise 18 (Green's function for the half-space). Let D denote the half-space $x_3 > 0$. The Green function of D is

$$G(\boldsymbol{x}|\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} - \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}'|} \tag{E7}$$

where $\boldsymbol{y}' = (y_1, y_2, -y_3)$ is the point, belonging to the complementary half-plane $x_3 < 0$, obtained from $\boldsymbol{y}= (y_1, y_2, y_3)$ by specular reflection with respect to the boundary plane $\partial D : x_3 = 0$. The potential $Q_{\rm o} g(\boldsymbol{x}|\boldsymbol{y})/\epsilon$ of the induced charges coincides with the Coulomb potential

$$\frac{-Q_{\rm o}}{4\pi\epsilon|\boldsymbol{x}-\boldsymbol{y}'|}$$

due to an "image charge" $Q'_o = -Q_o$ concentrated at the point $\boldsymbol{y}'$.

Exercise 19 (influence problem for a dielectric with isolated conducting boundary). Let $D = D_-$ denote a bounded dielectric with permittivity ϵ_- containing the influencing charge Q_o, and $D_+ = \mathbb{R}^3 \backslash \overline{D}_-$ another dielectric with permittivity ϵ_+. If the (connected and conducting) separation surface $\mathbb{S} = \partial D_-$ is isolated, its potential V is unknown and the total induced surface charge Q_i is zero, that is,

$$\text{(E8)} \qquad \int_{\mathbb{S}} \left[\epsilon_- \frac{\partial u_-(\boldsymbol{x})}{\partial n_x} - \epsilon_+ \frac{\partial u_+(\boldsymbol{x})}{\partial n_x} \right] dS_x \;=\; 0$$

The potential is given by

$$u_+(\boldsymbol{x}) = V u_1(\boldsymbol{x}) \;\;, \qquad u_-(\boldsymbol{x}) = \frac{Q_o}{\epsilon_-} G(\boldsymbol{x}|\boldsymbol{y}) + V \qquad (\boldsymbol{x} \neq \boldsymbol{y})$$

and since the Green function $G(\boldsymbol{x}|\boldsymbol{y})$ satisfies eq. (2.62), from eqs. (E8) and (2.46) we find

$$0 = \; Q_o \int_{\partial C} \frac{\partial G(\boldsymbol{x}|\boldsymbol{y})}{\partial n_x} \, dS_x - V\mathbb{C} \;=\; Q_o \; - V\mathbb{C}$$

or $V = Q_o/\mathbb{C}$, where $\mathbb{C}$ is the capacity of the separation surface $\mathbb{S}$.

Exercise 20. Find the electrostatic potential u due to a point charge Q_o concentrated at the origin in a spherical dielectric D_- of radius R_- surrounded by an isolated conducting spherical shell of radii R_-, R_+ and surrounded by another dielectric D_+ extending to infinity.

Hint. The solution is $u = V$ in the spherical shell $R_- < r < R_+$ and $u_{\mp}$ in the dielectrics, with

$$u_- = V + \frac{Q_o}{4\pi\epsilon_- r} - \frac{Q_o}{4\pi\epsilon_- R_-} \qquad , \qquad u_+ = \frac{V R_+}{r}$$

the induced surface charge densities are

$$\sigma_- = \frac{-Q_o}{4\pi R_-^2} \;\Rightarrow\; Q_- = -Q_o \quad ; \quad \sigma_+ = \frac{V}{R_+} \;\Rightarrow\; Q_+ = \mathbb{C}_+ V$$

where $\mathbb{C}_+ = 4\pi\epsilon_+ R_+$, and $Q_o = \mathbb{C}_+ V$.

Exercise 21 (capacity of the spherical condenser). The potential is radial for symmetry reasons and by the Gauss Law it coincides with the Coulomb potential $\frac{Q}{4\pi\epsilon r}$. It follows that

$$V_1 - V_2 = \frac{Q}{4\pi\epsilon}(\frac{1}{R_1} - \frac{1}{R_2})$$

and $\mathbb{C} = \frac{Q}{V_1 - V_2}$ is given by eq. (2.78).

Exercise 22. The electrostatic field $\boldsymbol{E}$ in a spherical condenser D is irrotational and solenoidal in D but no global vector potential exists in D.

Hint:Suppose $\boldsymbol{E}= curl\, \boldsymbol{V}$ in all of D. If $Q \neq 0$ $\boldsymbol{E}$ is different from zero and is radial for reasons of symmetry, so that $\boldsymbol{E}\wedge\boldsymbol{n}= \boldsymbol{0}$ on ∂D. We have then

$$\int_D |\boldsymbol{E}|^2 dV = \int_D curl\, \boldsymbol{V} \cdot \boldsymbol{E}\ dV = \int_D curl \boldsymbol{E} \cdot \boldsymbol{V} dV - \int_{\partial D} \boldsymbol{E} \wedge \boldsymbol{V} \cdot \boldsymbol{n} dS = 0$$

so that $\boldsymbol{E}\equiv \boldsymbol{0}$ in D, in contradiction with the previous Exercise.

Exercise 23. Proceeding as in Exercise 5 we find as $|\boldsymbol{x}|$ approaches infinity

$$\frac{M_o R^2}{4|\boldsymbol{x} - \boldsymbol{y}_2|} - \frac{M_o R^2}{4|\boldsymbol{x} - \boldsymbol{y}_1|} \sim \frac{2M_o R^2(\boldsymbol{y}_2 - \boldsymbol{y_1}) \cdot \boldsymbol{x}}{4|\boldsymbol{x}|^3} \sim l M_o R^2 \frac{\mathbf{c}_3 \cdot \boldsymbol{x}}{|\boldsymbol{x}|^3}$$

Exercise 24 (permanent magnet). Eq. (2.86) written for $x = y = 0$ yields

$$\begin{aligned} v(0,0,z) &= \frac{M_o}{4\pi} 2\pi \left\{ \int_0^R \frac{r\,dr}{\sqrt{r^2 + (l-z)^2}} - \int_0^R \frac{r\,dr}{\sqrt{r^2 + (l+z)^2}} \right\} \\ &= \frac{M_o}{2} \left\{ \sqrt{R^2 + (l-z)^2} - \sqrt{R^2 + (l+z)^2} \right\} + M_o z \end{aligned}$$

and eq. (2.86) is obtained by differentiation. Inside the magnet $l - z > 0$, $l + z > 0$ and therefore

$$H(z) = \frac{M_o}{2} \left\{ \frac{l-z}{\sqrt{(l-z)^2 + R^2}} + \frac{l+z}{\sqrt{(l+z)^2 + R^2}} - 2 \right\} \ < \ 0 \quad \text{for } |z| \leq l$$

Exercise 25. Show that the volume potential $\mathbb{V}$ due to a uniformly charged ball in eq. (2.14) coincides with the potential in Exercise 2.

Exercise 26. Show that the single layer potential $\mathcal{V}$ due to a uniformly charged sphere in eq. (2.16) is the same as in Exercise 1.

Exercise 27. Consider a simply connected bounded conductor C surrounded by a dielectric $D := \mathbb{R}^3 \backslash \overline{C}$ with permittivity ϵ. Let

$$\rho(\boldsymbol{x}) \equiv 0 \quad \text{in } D \ , \qquad \sigma(\boldsymbol{x}) \neq 0 \quad \text{on } \partial C$$

Find the unique electric field in D satisfying the asymptotic condition (2.5). Hint. $\boldsymbol{E} = -grad\, u$, with $u = V$ constant in C, and $u = V u_1$ with u_1 the capacitary potential of ∂C in D.

Chapter 3

Steady Currents In Conductors

In this chapter we examine the stationary electromagnetic field generated by steady currents in conductors, extending to arbitrary conductors and to solenoids the Biot-Savart law for direct currents in infinitely thin wires introduced in §1.1.5. We also debate the issue whether or not a stationary electric field $\boldsymbol{E}(\mathbf{x})$ exists outside a conductor carrying a steady current. From the mathematical point of view this involves essentially studying some properties of irrotational and solenoidal vector fields and finding the inverse of the *curl* operator in the space of solenoidal fields. The last section of this chapter is of a different nature and contains a brief account about the equations of time-dependent electric circuits in the quasi-stationary approximation.

Consider a single bounded, homogeneous and non-magnetic conductor C, surrounded by an unbounded homogeneous uncharged dielectric of permittivity ϵ extending to infinity in $\mathbb{R}^3$. We have then

$$\mu(\boldsymbol{x}) = \mu_o \ , \quad \rho(\boldsymbol{x}) \equiv 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^3$$

and the stationary magnetic field satisfies eqs.(1.58)

$$curl\, \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x}) \qquad , \qquad div\, \boldsymbol{H}(\boldsymbol{x}) = \mathbf{0}$$

where the stationary current density $\boldsymbol{J}(\boldsymbol{x}) = \gamma \boldsymbol{E}(\boldsymbol{x})$ in C cannot be given arbitrarily, but must be a solution of eqs. (1.59) and (1.60) :

(3.1) $$div\, \boldsymbol{J} = 0 \ , \ curl\, \boldsymbol{J} = 0 \ \text{ in } \ C \quad , \quad \boldsymbol{J} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial C$$

Thus $\boldsymbol{J}$ is an irrotational and solenoidal vector field in C with zero normal trace on ∂C. If C is contourwise simply connected, Lemma 1.6.2, Counterexample 1 and Corollary 1.6.10 show that $\boldsymbol{J} \equiv \mathbf{0}$ and then, if the power flux at infinity is zero, also $\boldsymbol{H} \equiv \mathbf{0}$: in order to have a non-vanishing steady current (and magnetic field) we must assume that the conductor C is contourwise multiply connected.

3.1 Neumann vector fields.

The simplest case of a contourwise multiply connected (c.m.c.) domain is that of a single toroidal [1] conductor $\mathcal{T}$ with smooth boundary $\partial\mathcal{T}$. Suppose that $\boldsymbol{J}(\boldsymbol{x}) \in C^1(\overline{\mathcal{T}})$. By force of eqs. (3.1) we can write

$$\boldsymbol{J} = -\gamma \, grad \, u$$

where the interior potential $u(\boldsymbol{x})$ satisfies the homogeneous Neumann problem for $\mathcal{T}$

$$(3.2) \qquad \Delta_3 u = 0 \quad \text{for } \boldsymbol{x} \in \mathcal{T} \quad , \qquad \frac{\partial u}{\partial n} = 0 \quad \text{for } \boldsymbol{x} \in \partial\mathcal{T}$$

For a c.s.c. domain, the potential u would be one-valued, the solution u would be constant, and so $\boldsymbol{J}$ would be identically zero. Besides, closed current lines would be impossible (see Exercise 4 of Chapter 1). In contrast, a toroidal domain such as $\mathcal{T}$ admits closed current lines and has the following important topological property:

There exists two classes of closed curves $\Gamma \subset \mathcal{T}$, $\Gamma' \subset \mathcal{T}' = \mathbb{R}^3 \backslash \overline{\mathcal{T}}$ that are *irreducible*, in the sense that they cannot be reduced by continuous deformation to a point without crossing $\partial\mathcal{T}$. Correspondingly, there exists two classes of closed curves Γ^*, $\Gamma'^* \subset \partial\mathcal{T}$, called homology classes of curves on $\partial\mathcal{T}$ [11, 25], such that a curve of one class cannot be reduced by continuous deformation into one of the other class.

The circulation of $\boldsymbol{J}/\gamma$ along any closed irreducible path Γ is no longer required to be zero and defines the period p of the many-valued function u

[1] this means that $\partial\mathcal{T}$ is a surface of "topological genus $p = 1$", homeomorphic to a sphere with one handle

in $\mathcal{T}$:

$$\oint_\Gamma \boldsymbol{J} \cdot \boldsymbol{t}\, ds = -\gamma p \tag{3.3}$$

It is well-known that the period p does not depend on the choice of the particular irreducible curve Γ (Exercise 1). If $p \neq 0$, the potential $u(\boldsymbol{x})$ is a many-valued solution of (3.2) in $\mathcal{T}$ and is no longer constant, so that the corresponding vector field

$$\boldsymbol{N}(\boldsymbol{x}) := -grad\, u \qquad (\boldsymbol{x} \in \mathcal{T})$$

called a Neumann vector field for $\mathcal{T}$, no longer vanishes identically. $\boldsymbol{N}(\boldsymbol{x})$ and u are defined up to an arbitrary non-null factor which is fixed by fixing the period.

In an entirely similar way one can define a Neumann vector field

$$\boldsymbol{N}'(\boldsymbol{x}) := -grad\, u' \qquad (\boldsymbol{x} \in \mathcal{T}')$$

for the c.m.c. exterior domain $\mathcal{T}' = \mathbb{R}^3 \backslash \overline{\mathcal{T}}$. We will denote by a prime the quantities pertaining to $\mathcal{T}'$.

Proposition 3.1.1 *The Neumann vector field* $\boldsymbol{N}(\boldsymbol{x})$*=-grad* u *is a one−valued* C^∞ *irrotational and solenoidal vector function in* $\mathcal{T}$*, determined up to an arbitrary factor defined by the circulation along an irreducible curve* Γ

$$\oint_\Gamma \boldsymbol{N} \cdot \boldsymbol{t}\, ds = -p$$

Moreover, $\boldsymbol{N} \in C^1(\overline{\mathcal{T}})$ *and satisfies the boundary conditions on* $\partial\mathcal{T}$

$$\boldsymbol{N} \cdot \boldsymbol{n} = 0 \ \ \text{on } \partial\mathcal{T}, \quad \boldsymbol{N} \wedge \boldsymbol{n} \text{ is not identically zero on } \partial\mathcal{T}\,, \quad \int_{\partial\mathcal{T}} \boldsymbol{N} \wedge \boldsymbol{n}\, dS = \boldsymbol{0}$$

Proof. Since the different branches of the many-valued potential u differ additively by multiples of the period p, $grad u$ is one-valued in $\mathcal{T}$. Thus the irrotational and solenoidal vector field $\boldsymbol{N}(\boldsymbol{x})$ is one-valued and, by the vector version of the Weyl Lemma [2], is of class $C^\infty(\mathcal{T})$. Because $\mathcal{T}$ is

[2] see H. Weyl, "The method of orthogonal projections in potential theory", Duke Math. J. 7, 1940, p. 411

s.s.c., Lemma 1.6.1 implies that $\boldsymbol{N}\wedge\boldsymbol{n}$ cannot be identically zero on $\partial\mathcal{T}$. On the other hand, the integral of $\boldsymbol{N}\wedge\boldsymbol{n}$ on $\partial\mathcal{T}$ (called vector circulation in aerodynamics) must vanish because of the Gauss Lemma

$$\int_{\partial T}\boldsymbol{N}\wedge\boldsymbol{n}\,dS = \int_{T} curl\boldsymbol{N}\ dV = \mathbf{0}$$

The electric current $\boldsymbol{J}(\boldsymbol{x})$ is a Neumann vector field in $\mathcal{T}$. So is $\boldsymbol{E}= \gamma^{-1}\boldsymbol{J}$ in a toroidal conductor for $0 < \gamma < +\infty$.

Example (Neumann vector fields for the interior of a torus). Consider a homogeneous c.m.c. conductor $\mathcal{T}$ bounded by a torus $\partial\mathcal{T} = \mathbb{C}_a \times \mathbb{C}_b$,where $\mathbb{C}_a$ is a circumference of radius a centered at the origin and lying in the $(x, y)-$plane and $\mathbb{C}_b$ is a circumference of radius $b < a$ centered at the point $(a, 0, 0)$ and lying in the $(x, z)-$plane (see Fig. 3.1).

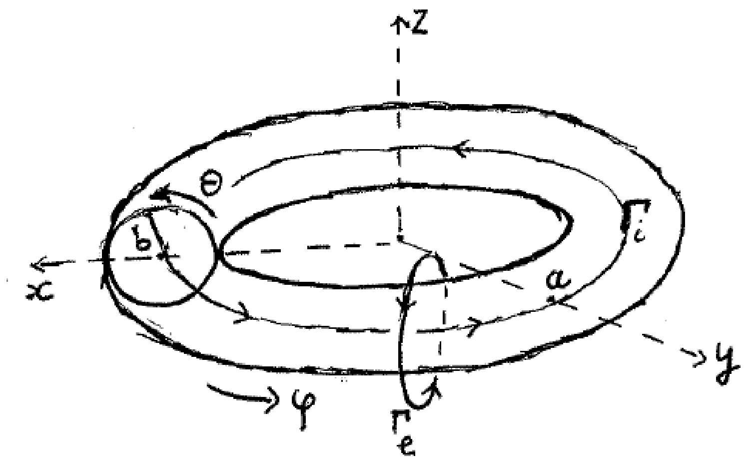

Figure 3.1: Irreducible curves $\Gamma = \Gamma_i$, $\Gamma' = \Gamma_e$ for the torus

A Neumann field for the interior $\mathcal{T}$ of the torus is given by

$$\boldsymbol{N}\,(\boldsymbol{x}) = \frac{M}{x^2 + y^2}(x\mathbf{j} - y\mathbf{i})$$

where $M \neq 0$ is an arbitrary factor, $\mathbf{x}= (x, y, z)$, and $\mathbf{i}$, $\mathbf{j}$ are the unit vectors of the x and y axes, respectively. If

$$\varrho := \sqrt{x^2 + y^2} \quad , \quad \varphi := arctan(y/x)$$

it is immediate to check that $div\,\boldsymbol{N}\,(\boldsymbol{x}) = curl\,\boldsymbol{N}\,(\boldsymbol{x}) = 0$ for $\varrho > 0$, $0 \leq \varphi \leq 2\pi$. Let θ $(0 \leq \theta \leq 2\pi)$ denote the coordinate angle in each cross section of $\mathcal{T}$ and r $(0 \leq r \leq b)$ the radial distance of $\boldsymbol{x}$ from the centerline of the sections , so that

$$\varrho = \ a - r cos\theta$$

The equations of $\mathcal{T}$ in the system of coordinates (r, θ, φ) adapted to the torus are then

$$\mathcal{T} : \qquad x = (a - rcos\theta)cos\varphi \ , \ \ y = (a - rcos\theta)sin\varphi \, , \ \ z = rsin\theta$$

and the equations of the torus $\partial\mathcal{T}$ are obtained by setting $r = b$. In this system of coordinates we have

$$\boldsymbol{N}\,(\boldsymbol{x}) \ = \ \frac{M}{a - rcos\theta}(cos\varphi\,\mathbf{j} - sin\varphi\,\mathbf{i}) \ \equiv -M grad\,u$$

where the inner potential

$$u(\boldsymbol{x}) := -M\varphi \qquad\qquad \text{for} \ \ \boldsymbol{x} \in \mathcal{T}$$

is harmonic and many-valued in $\mathcal{T}$, with period

$$p \ = -2\pi M$$

Thus all orthogonal cross sections φ =constant are equipotential The lines of current Γ of $\boldsymbol{N}$ are φ-lines

$$\Gamma \ \ : \qquad r = \text{constant} \ \ , \quad \theta = \text{constant}$$

i.e. circumferences of constant radius ϱ in the (x, y)-plane. Hence all the current lines Γ are closed and constitute a class of irreducible curves for $\mathcal{T} \cup \partial\mathcal{T}$. Obviously $\boldsymbol{N} \cdot \boldsymbol{n} = 0$ on $\partial\mathcal{T}$, as required, and the circulation is

$$\oint_\Gamma \boldsymbol{N} \ \cdot \boldsymbol{t}\, ds = -p \ \equiv \ 2\pi M \ \neq \ 0$$

The vector field $\boldsymbol{N}\,(\boldsymbol{x})$ is determined up to the multiplicative factor M to which p is proportional, and the total current through any section Σ of $\mathcal{T}$

$$I = \int_\Sigma \ \boldsymbol{J} \cdot \boldsymbol{n} \ \ dS$$

is proportional to the period:

$$I = M \int_0^b r dr \int_0^{2\pi} \frac{d\theta}{a - r cos\theta} = \kappa M = -\frac{\kappa}{2\pi} p \neq 0 \tag{3.4}$$

where κ is the result of the calculation of the double integral, namely

$$\kappa = 2\pi(a - \sqrt{a^2 - b^2})$$

$\boldsymbol{N}(\boldsymbol{x})$ is bounded and nowhere zero in $\mathcal{T} \cup \partial\mathcal{T}$, but it is easy to see that the vector circulation of $\boldsymbol{N}$ on $\partial\mathcal{T}$ vanishes. The Neumann field $\boldsymbol{N}'$ for $\mathcal{T}'$ is of the form

$$\boldsymbol{N}'(\boldsymbol{x}) = F(r, \theta)\, sin\theta\, (sin\varphi\, \mathbf{j} + cos\varphi \mathbf{i}) \equiv -grad\, u'$$

and cannot be determined in terms of elementary functions for $\mathbf{x} \in \mathcal{T}'$, whereas for $\boldsymbol{x} \in \partial\mathcal{T}$ the exterior potential is given by

$$u' = \frac{1}{\pi} arctan\Big[\sqrt{\frac{a-b}{a+b}}\, tan\frac{\theta}{2}\Big] + \text{const.} \qquad (\boldsymbol{x} \in \partial\mathcal{T})$$

Thus $\boldsymbol{N}'(\boldsymbol{x})$ is parallel to the θ-lines

$$\Gamma' : \quad r = b,\ \varphi = \text{constant}$$

which constitute a second class of irreducible curves on $\partial\mathcal{T}$.

In the preceding example $\boldsymbol{N}$ and $\boldsymbol{N}'$ are orthogonal and hence cannot match continuously on $\partial\mathcal{T}$. This is true in general.

Proposition 3.1.2 *Suppose the Neumann field $\boldsymbol{N}'(\boldsymbol{x})$ for the exterior domain $\mathcal{T}'$ satisfies the uniform asymptotic condition at infinity*

$$\boldsymbol{N}'(\boldsymbol{x}) = O(|\boldsymbol{x}|^{-2}) \qquad as\ |\boldsymbol{x}| \to \infty \tag{3.5}$$

Then $\boldsymbol{N}'(\boldsymbol{x}) = -grad\, u'$ cannot match continuously on $\partial\mathcal{T}$ with the Neumann vector field $\boldsymbol{N}(\boldsymbol{x}) = -grad\, u$ for the interior domain $\mathcal{T}$ unless their periods are zero:

$$p = p' = 0$$

so that u, u' are one-valued and $\boldsymbol{N}(\boldsymbol{x}) \equiv \boldsymbol{N}'(\boldsymbol{x}) \equiv \mathbf{0}$

Proof. Suppose $\boldsymbol{N} = \boldsymbol{N}'$ on $\partial\mathcal{T}$. Then the circulation of $\boldsymbol{N}$ along any irreducible Γ curve on $\partial\mathcal{T}$ is given by p, and the circulation of $\mathrm{N}' = \boldsymbol{N}$ along any irreducible Γ' curve on $\partial\mathcal{T}$ is given by p'. As $\mathcal{T}$ and $\mathcal{T}'$ are s.s.c. we can construct a surface $S \subset \mathcal{T}$ and a surface $S' \subset \mathcal{T}'$ such that $\Gamma = \partial S'$ and $\Gamma' = \partial S$. By applying Stokes' theorem (ST1) of Chapter I we obtain

$$(3.6) \qquad \begin{aligned} p &= -\oint_{\partial S'} \boldsymbol{N}' \cdot \boldsymbol{t}\, ds = -\int_{S'} curl\boldsymbol{N}' \cdot \boldsymbol{n}\, dS = 0 \\ p' &= -\oint_{\partial S} \boldsymbol{N} \cdot \boldsymbol{t}\, ds = -\int_{S} curl\boldsymbol{N} \cdot \boldsymbol{n}\, dS = 0 \end{aligned}$$

Hence u and u' are one-valued, and all the hypotheses of Corollary 1.6.4 are satisfied. It follows that u is constant in $\mathcal{T}$, u' is constant in $\mathcal{T}'$, and $\boldsymbol{N} \equiv \boldsymbol{N}' \equiv \boldsymbol{0}$.

We now show that if $\boldsymbol{J}(\boldsymbol{x})$ is not identically zero in $\mathcal{T}$ a paradox arises. Consider the electric field $\boldsymbol{E}'$ in the dielectric $\mathcal{T}'$. The stationary Maxwell equations with $\rho \equiv 0$ say that (see Chapter 1)

$$(3.7) \qquad curl\boldsymbol{E}'(\boldsymbol{x}) = \boldsymbol{0} \quad , \quad div\, \boldsymbol{E}'(\boldsymbol{x}) = 0 \qquad \text{for } \boldsymbol{x} \in \mathcal{T}'$$

$$(3.8) \qquad \boldsymbol{E}_+ \cdot \boldsymbol{n} = \frac{\sigma}{\epsilon} \qquad for\ \boldsymbol{x} \in \partial\mathcal{T}$$

where σ is the surface charge density on $\partial\mathcal{T}$, and $\boldsymbol{E}_+$ is the trace of $\boldsymbol{E}'$ on $\partial\mathcal{T}$. The boundary condition (3.8) can be obtained from the constitutive relation $\boldsymbol{D}= \epsilon\boldsymbol{E}$, eq. (3.1) and the matching relation (R1) (Chapter 1) across $\partial\mathcal{T}$:

$$\boldsymbol{E}_- \cdot \boldsymbol{n} = \frac{1}{\gamma}\boldsymbol{J} \cdot \boldsymbol{n} = 0 \;\Rightarrow\; \boldsymbol{D}_- \cdot \boldsymbol{n} = 0 \quad \Rightarrow\; \epsilon\boldsymbol{E}_+ \cdot \boldsymbol{n} = \boldsymbol{D}_+ \cdot \boldsymbol{n} = \sigma$$

where $\boldsymbol{E}_-$ denotes the trace of $\boldsymbol{E}$. Moreover, we know from (R5) of Chapter 1 that $\boldsymbol{E}\wedge\boldsymbol{n}$ is continuous across $\partial\mathcal{T}$.

Corollary 3.1.3 *If $\boldsymbol{J}(\boldsymbol{x}) \neq \boldsymbol{0}$ in a toroidal conductor $\mathcal{T}$, then $\boldsymbol{E}'(\boldsymbol{x})$ cannot be irrotational in the dielectric $\mathcal{T}'$.*

Proof. By assumption, $\boldsymbol{E}= \gamma^{-1}\boldsymbol{J}$ is a non-zero Neumann vector field in $\mathcal{T}$. Moreover the tangential component of $\boldsymbol{E}$ is continuous across $\partial\mathcal{T}$.

Thus if Γ^* is any irreducible curve of the first (homology) class on $\partial\mathcal{T}$, the circulation of $\boldsymbol{E} = -grad\, u$

$$\oint_{\Gamma^*} \boldsymbol{E} \cdot \boldsymbol{t}\, ds = -p \tag{3.9}$$

does not vanish, and by applying Stokes' theorem as in the proof of Proposition 3.1.2 we find that $curl\, \boldsymbol{E}'$ cannot vanish identically in $\mathcal{T}'$.

This corollary is a stronger version of Theorems 1.6.5 and 1.6.6 and shows that in steady conditions $\boldsymbol{E}$ and $\boldsymbol{J}$ must be identically zero in the conductor: a conductor can never support a steady current, even if it is toroidal. In pont of fact, if $\boldsymbol{J}(\boldsymbol{x}) \neq \boldsymbol{0}$ in a toroidal conductor the circulation of $\boldsymbol{E}$ along any curve Γ must vanish (see the proof of the above corollary) while the circulation of $\boldsymbol{J}$ must remain different from zero, and this is incompatible with the homogeneous Maxwell equations considered so far.

A steady electric current in a closed conductor is possible only if the conductor is toroidal and if the Joule dissipation is balanced by an e.m.f. generator. The latter can be modeled, following Heaviside, by introducing a suitable "impressed electric field" which appears as a source term in the Maxwell equations.

3.2 Inhomogeneous Maxwell equations.

3.2.1 Impressed electric field

The Maxwell equations (M1)−(M5) considered in §1.3.1 of Chapter 1 are not suitable to describe all relevant electromagnetic phenomena, as first remarked by Heaviside [30]. In many instances one must take into account phenomena of non-electromagnetic (chemical, thermal, mechanical...) origin, and this can be done by introducing an assigned *impressed electric field* $\boldsymbol{E}_{\text{imp}}$. A typical example is an electric circuit with an e.m.f. generator, as anticipated in the previous section. In the case of a conductor with conductivity γ $(0 < \gamma < \infty)$, an impressed electric current $\boldsymbol{J}_{\text{imp}}$ may also be defined by

$$\boldsymbol{J}_{\text{imp}}(\boldsymbol{x}, t) := \gamma \boldsymbol{E}_{\text{imp}}(\boldsymbol{x}, t) \tag{3.10}$$

and Ohm's law $\boldsymbol{J} = \gamma\boldsymbol{E}$ must be replaced by

$$\boldsymbol{J} = \gamma\,(\boldsymbol{E} + \boldsymbol{E}_{\text{imp}}) \equiv \gamma\boldsymbol{E} + \boldsymbol{J}_{\text{imp}} \tag{3.11}$$

In the case of a linear current along a thin wire $\Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2)$ where $\boldsymbol{E}_{\mathrm{imp}} \neq \boldsymbol{0}$, the modified Ohm's law (3.11) takes the form

$$\mathcal{R}I = V + e_{\mathrm{i}} \tag{3.12}$$

where V is the potential drop of $\boldsymbol{E}$ and e_{i} is the impressed e.m.f.

$$e_{\mathrm{i}} = \int_{\Gamma(\boldsymbol{x}_1,\boldsymbol{x}_2)} \boldsymbol{E}_{\mathrm{imp}} \cdot \boldsymbol{t}\, ds$$

The Faraday induction law takes the form

$$\frac{d\Phi}{dt} = e_{\mathrm{i}} - \mathcal{R}I \tag{3.13}$$

where Φ is the magnetic flux and I is the induced current. Since by eq. (3.12) $e_{\mathrm{i}} = \mathcal{R}I - V$, this is the same equation

$$\frac{d\Phi}{dt} = -V$$

as before (eq. (1.22)). Hence the Faraday induction law needs no modification in the presence of an impressed field. Similarly the displacement vector $\boldsymbol{D}$ is independent of the impressed field. The general (time-dependent) inhomogeneous Maxwell equations are then

$$\begin{gathered} \frac{\partial \boldsymbol{D}}{\partial t} - curl\, \boldsymbol{H} + \gamma \boldsymbol{E} = -\boldsymbol{J}_{\mathrm{imp}}(\boldsymbol{x},t), \quad \frac{\partial \boldsymbol{B}}{\partial t} + curl\, \boldsymbol{E} = 0 \\ div\, \boldsymbol{D} = \rho, \quad div\, \boldsymbol{B} = 0 \end{gathered} \tag{3.14}$$

and the impressed current $\boldsymbol{J}_{\mathrm{imp}}(\boldsymbol{x},t) = \gamma \boldsymbol{E}_{\mathrm{imp}}(\boldsymbol{x},t)$ appears as a known source term. For homogeneous non-magnetic media, the linear constitutive relations

$$\boldsymbol{D} = \epsilon \boldsymbol{E}, \;\; \boldsymbol{B} = \mu_o \boldsymbol{H}$$

remain unchanged and eqs. (3.14) become

$$\begin{aligned} &\epsilon \frac{\partial \boldsymbol{E}}{\partial t} - curl\, \boldsymbol{H} + \gamma \boldsymbol{E} = -\boldsymbol{J}_{\mathrm{imp}}(\boldsymbol{x},t), \quad \mu_o \frac{\partial \boldsymbol{H}}{\partial t} + curl\, \boldsymbol{E} = 0 \\ &div \boldsymbol{E} = \rho/\epsilon\,, \quad div \boldsymbol{H} = 0 \end{aligned}$$

In a conductor Ω we have $\rho \equiv 0$ and the impressed current satisfies

$$div\, \boldsymbol{J}_{\text{imp}}(\boldsymbol{x}, t) = 0 \;\text{ in } \Omega \quad , \quad \mathbf{J}_{\text{imp}} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \tag{3.15}$$

Let S denote the (given) support of $\boldsymbol{E}_{\text{imp}}(\boldsymbol{x}, t)$, i.e. of $\boldsymbol{J}_{\text{imp}}(\boldsymbol{x}, t)$. By proceeding as in §1.5 we easily obtain the modified form of the energy balance equation for a domain Ω of $\mathbb{R}^3$

$$\frac{d\mathcal{E}[\Omega]}{dt} = -\int_\Omega \gamma |\boldsymbol{E}|^2\, \boldsymbol{dx} \; - \int_{\partial\Omega} \boldsymbol{S} \cdot \boldsymbol{n}\, dS \;\; + P_{\text{imp}}[\Omega]$$

where

$$P_{\text{imp}}[\Omega] := \begin{cases} -\int_\Omega \boldsymbol{J}_{\text{imp}} \cdot \boldsymbol{E}\, \boldsymbol{dx} & \text{if} \quad S \subseteq \Omega \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

For a conductor in steady conditions the inhomogeneous Maxwell equations (3.14) become

$$curl\, \boldsymbol{H} = \boldsymbol{J}(\boldsymbol{x}) \quad , \quad curl\, \boldsymbol{E} = 0 \quad , \quad div\, \boldsymbol{E} \; = div\, \boldsymbol{H} \; = 0$$

where $\boldsymbol{J}$ satisfies (3.11), and the energy balance equation takes the form

$$\int_\Omega \gamma |\boldsymbol{E}|^2\, \boldsymbol{dx} + \int_{\partial\Omega} \boldsymbol{S} \cdot \boldsymbol{n}\, dS \;\; = \;\; P_{\text{imp}}[\Omega]$$

where P_{imp} is defined by (3.16). Moreover, eqs. (3.15) imply that $\boldsymbol{J}(\boldsymbol{x})$ satisfies

$$div\, \boldsymbol{J}(\boldsymbol{x}) = 0 \;\text{ in } \Omega, \quad \boldsymbol{J} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \tag{3.17}$$

Eq. (3.11) can be extended to quasi-stationary fields and in particular to an a-c electric circuit, briefly considered at the end of this Chapter.

An impressed magnetic field has also been introduced by Heaviside, but we will never encounter situations where this concept is needed .

3.2.2 Potential jump model for toroidal conductors.

In applying these ideas to the toroidal conductor $\mathcal{T}$ in steady conditions, we need the following assumptions.

H1. (i) The level surfaces $u(\boldsymbol{x})$ =constant are sections Σ of $\mathcal{T}$

(ii) The total current $I \neq 0$, and $I > 0$ by orienting the normals $\boldsymbol{n}$ to every level surface in the same direction as $\boldsymbol{J}$.

These assumptions are plausible from a physical point of view and are satisfied in the case of small perturbations of a torus.

The impressed electric field $\boldsymbol{E}_{\rm imp}(\boldsymbol{x})$ must be such that the electric field $\boldsymbol{E}(\boldsymbol{x})$ has zero circulation, while $\boldsymbol{J}= -\gamma grad\, u(\boldsymbol{x})$ is a Neumann field satisfying (3.11) and (3.17) for $\Omega = \mathcal{T}$. We can fulfill these requirements by taking

$$\boldsymbol{J}(\boldsymbol{x}) = -\gamma grad\, u(\boldsymbol{x}) \quad , \quad \boldsymbol{E}\,(\boldsymbol{x}) = -grad\, \tilde{u}(\boldsymbol{x}) \qquad \boldsymbol{x} \in \mathcal{T} \tag{3.18}$$

where $u(\boldsymbol{x})$ is a many-valued potential of period $p < 0$ satisfying (3.2), and $\tilde{u}(\boldsymbol{x})$ is a one-valued branch of $u(\boldsymbol{x})$ obtained by introducing a *branch surface* Σ_d for u. In this way Σ_d is a level surface of u, $\tilde{u}(\boldsymbol{x})$ is discontinuous through Σ_d, and the discontinuity jump $e_{\rm i} > 0$ of $\tilde{u}(\mathbf{x})$ is equal and opposite to the period p of the many-valued potential u in $\mathcal{T}$:

$$e_{\rm i} = -p \tag{3.19}$$

This potential jump $e_{\rm i}$ represents the impressed e.m.f., and the branch surface Σ_d represents the generator, the two sides of Σ_d representing the two electrodes [27]. As soon as $e_{\rm i} > 0$ is assigned, $p < 0$ is known from eq. (3.19) and both u and $\tilde{u}$ are known.

The electric field in the conductor can also be written in the form

$$\boldsymbol{E}(\boldsymbol{x}) = \ - grad\, u(\boldsymbol{x}) - \boldsymbol{E}_{\rm imp}(\boldsymbol{x}) \ \equiv \ \gamma^{-1}\boldsymbol{J} - \boldsymbol{E}_{\rm imp} \tag{3.20}$$

where the impressed field $\boldsymbol{E}_{\rm imp}$ is a Dirac $\delta-$distributions with support Σ_d :

$$\boldsymbol{E}_{\rm imp} = \ e_{\rm i}\delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}) \tag{3.21}$$

($\delta = \delta_{\Sigma_d}$, $\boldsymbol{n}(\boldsymbol{x})$ the normal to Σ_d at the point $\boldsymbol{x}$). In this way $\boldsymbol{J}= \gamma\,(\boldsymbol{E}+\boldsymbol{E}_{\rm imp})$ satisfies (3.11). The equation $curl\,\boldsymbol{E}= \mathbf{0}$ is satisfied outside the generator, i.e. in $\mathcal{T}\backslash\Sigma_d$, and the e.m.f. $e_{\rm i}$ equals the circulation of $\boldsymbol{E}_{\rm imp}$ along any current line Γ of $\boldsymbol{J}$

$$\oint_\Gamma \boldsymbol{E}_{\rm imp} \cdot \boldsymbol{t}\, ds \equiv \ e_{\rm i}$$

so that by eq. (3.19) we have

$$\oint_\Gamma \boldsymbol{E} \cdot \boldsymbol{t}\, ds = -\oint_\Gamma grad\, u(\boldsymbol{x}) \cdot \boldsymbol{t}\, ds - \oint_\Gamma \boldsymbol{E}_{\rm imp} \cdot \boldsymbol{t}\, ds = -p \;-\; e_{\rm i} \;= 0$$

Thus the circulation of $\boldsymbol{E}$ vanishes, as required, whereas the current $\boldsymbol{J} = -\gamma grad\, u$ in the conductor remains unaltered and is still given by the Neumann vector field $\boldsymbol{N}$ with the same non-zero circulation $-\gamma p > 0$.

To summarize: In order to have a non-vanishing steady current the conductor must be toroidal, and an e.m.f. generator must be added to the Maxwell equations in the form of an impressed electric field obtained by introducing a branch surface for the potential of $\boldsymbol{E}$, whereas the potential of $\boldsymbol{J}$ remains many-valued.

The case of several toroidal conductors, treated in §3.4, is similar.

3.3 Single toroidal conductor

3.3.1 The magnetic field: Biot-Savart law.

Consider a single homogeneous non-magnetic toroidal conductor $\mathcal{T}$ of conductivity $0 < \gamma < +\infty$ (like a copper wire) carrying a steady current $\boldsymbol{J}(\boldsymbol{x})$. In the simple potential jump model chosen in the previous section, Ohm's law $\boldsymbol{J} = \gamma \boldsymbol{E}$ holds almost everywhere in $\mathcal{T}$, since one must exclude only a two-dimensional section Σ_d (where $\boldsymbol{E}_{\rm imp} \neq \boldsymbol{0}$) and the distribution of the current inside the conductor is unaffected by the presence of the branch surface Σ_d. $\boldsymbol{J}(\boldsymbol{x})$ is thus a Neumann vector field $\boldsymbol{N}(\boldsymbol{x})$ in $\mathcal{T}$, determined up to a factor which is fixed by fixing the impressed e.m.f. $e_{\rm i}$.

The total steady current I in the conductor is given by the integral

$$I = \int_\Sigma \boldsymbol{J} \cdot \boldsymbol{n} \;\; dS$$

over any section Σ of the conductor $\mathcal{T}$.

Proposition 3.3.1 *The total steady current I is independent of the particular section Σ.*

Proof. Let Ω be a portion of the conductor $\mathcal{T}$ between two arbitrary sections Σ', Σ'' having normals $\mathbf{n}', \boldsymbol{n}''$ with the same (arbitrary) orientation. From eqs. (3.1) and the divergence theorem we have

$$0 = \int_{\Omega} div\, \boldsymbol{J}\, \boldsymbol{dx} = \int_{\partial\Omega} \boldsymbol{J} \cdot \boldsymbol{n}\, dS = \int_{\Sigma'} \boldsymbol{J} \cdot \boldsymbol{n}'\, dS - \int_{\Sigma''} \boldsymbol{J} \cdot \boldsymbol{n}''\, dS$$

This proves that the above integral defining I is independent of Σ. Under assumption **H1** of the previous section we can choose equipotential sections Σ, Σ', Σ'' and orient them so that $I > 0$. The total current I is then known by fixing the impressed e.m.f. e_{i} and conversely, e_{i} is known by fixing the total current I.

In other words, the current density $\boldsymbol{J}(\boldsymbol{x})$ is proportional to I and is completely determined by fixing either the current I or the impressed e.m.f. e_{i} in the conductor $\mathcal{T}$. In order to find the magnetic field generated by $\boldsymbol{J}$, we need to solve the stationary Maxwell equations

$$(3.22) \qquad curl\, \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x}) \;, \qquad div\, \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{0}$$

for $\boldsymbol{x} \in \mathbb{R}^3 \backslash \partial\mathcal{T}$, that is to determine the inverse of the *curl* operator

$$\boldsymbol{H} = curl^{-1} \boldsymbol{J}$$

in the space of solenoidal vectors, with $\boldsymbol{J} \in C^1(\overline{\mathcal{T}})$. If the inverse operator $curl^{-1}$ exists, the resulting magnetic field will be proportional to the total current I, and we will obtain a unique solution for the normalized magnetic field $\boldsymbol{H}(\boldsymbol{x})/I$.

Note that $\boldsymbol{J}(\boldsymbol{x})$ is identically zero in $\mathcal{T}' = \mathbb{R}^3 \backslash \overline{\mathcal{T}}$ but is different from zero on $\partial\mathcal{T}$ (see Proposition 3.1.1), so that $curl\, \boldsymbol{H} = \boldsymbol{J}$ has a jump discontinuity across $\partial\mathcal{T}$. On the other hand, since $\mu \equiv \mu_o$ everywhere, the matching relations (R2) and (R4) of Chapter 1 imply that $\boldsymbol{H}$ is continuous in $\mathbb{R}^3$.

Proposition 3.3.2 *Suppose that $\boldsymbol{H}(\boldsymbol{x})$ satisfies the asymptotic condition at infinity (3.5). Then the inverse operator $curl^{-1}$ exists and is represented by the Biot-Savart formula for bulk conductors*

$$(3.23) \qquad \boldsymbol{H}(\boldsymbol{x}) = curl\; \boldsymbol{V}(\boldsymbol{x}) \equiv \frac{1}{4\pi} \int_{\mathcal{T}} \boldsymbol{J}(\boldsymbol{y}) \wedge \frac{\boldsymbol{x} - \boldsymbol{y}}{|\boldsymbol{x} - \boldsymbol{y}|^{\,3}}\, \boldsymbol{dy} \;, \quad \boldsymbol{x} \in \mathbb{R}^3$$

where the vector potential $\boldsymbol{V}(\boldsymbol{x})$*, defined by*

$$\boldsymbol{V}(\boldsymbol{x}) := \frac{1}{4\pi}\int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\,\boldsymbol{dy} \equiv \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{\boldsymbol{J}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\,\boldsymbol{dy} \tag{3.24}$$

is solenoidal in $\mathbb{R}^3$*.*

Proof. By Lemma 1.6.1 the solution $\boldsymbol{H}$, if it exists, is unique. Since $\boldsymbol{J}\in C^1(\overline{\mathcal{T}})$, the properties of the volume potential (§2.1) imply that V$\in C^1(\mathbb{R}^3)\cap C^2(\mathbb{R}^3\backslash\partial\mathcal{T})$, so that eq. (3.23) defines a vector field $\boldsymbol{H}\in C^{o}(\mathbb{R}^3)\cap C^1(\mathbb{R}^3\backslash\partial\mathcal{T})$. It remains only to check that (3.23) yields an actual solution. As $\mathcal{T}$ is bounded, $\boldsymbol{H}$ satisfies the asymptotic condition at infinity (3.5) (Exercise 2), and clearly *div* $\boldsymbol{H}=0$. Taking the *curl* of $\boldsymbol{H}$ for $\boldsymbol{x}\notin\partial\mathcal{T}$ yields

$$curl_x \boldsymbol{H}(\boldsymbol{x}) = curl\, curl_{\mathbf{x}} \int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{y})}{4\pi r}\boldsymbol{dy} \tag{3.25}$$

$$\equiv grad\, div_{\mathbf{x}} \int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{y})}{4\pi r}\boldsymbol{dy} - \Delta_3 \int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{y})}{4\pi r}\boldsymbol{dy}$$

and so

$$curl_{\mathbf{x}}\, \mathbf{H}(\mathbf{x}) \equiv\ grad\, div_{\mathbf{x}}\mathbf{V}(\mathbf{x}) + \mathbf{J}(\mathbf{x}) \tag{3.26}$$

where $r = |\boldsymbol{x}-\boldsymbol{y}|$, and $\boldsymbol{J}(\boldsymbol{x}) \equiv \mathbf{0}$ for $\boldsymbol{x}\in \mathbb{R}^3\backslash\overline{\mathcal{T}}$. We have used here the identity (1.76) and Theorem 2.1.1. In order to obtain the desired result it remains to prove that $\boldsymbol{V}$ is solenoidal: indeed, by means of formal manipulations, using eq. (3.1) and the divergence theorem, we find that the divergence of $\boldsymbol{V}$ is given by

$$div_{\mathbf{x}} \int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{y})}{4\pi r}\boldsymbol{dy} = \frac{1}{4\pi}\int_{\mathcal{T}} \boldsymbol{J}(\boldsymbol{y})\cdot grad_{\mathbf{x}}\frac{1}{r}\boldsymbol{dy} \tag{3.27}$$

$$\equiv -\frac{1}{4\pi}\int_{\mathcal{T}} \boldsymbol{J}(\boldsymbol{y})\cdot grad_{\mathbf{y}}\frac{1}{r}\boldsymbol{dy}$$

$$= -\frac{1}{4\pi}\int_{\mathcal{T}} div_{\mathbf{y}}(\boldsymbol{J}(\boldsymbol{y})\frac{1}{r})\,\boldsymbol{dy} + \frac{1}{4\pi}\int_{\mathcal{T}} \frac{1}{r}\ div_{\mathbf{y}}\,\boldsymbol{J}(\boldsymbol{y})\boldsymbol{dy}\ = -\frac{1}{4\pi}\int_{\partial\mathcal{T}} \boldsymbol{n}\cdot\boldsymbol{J}\frac{1}{r}\ dS = 0$$

The result of this formal calculation is correct for $\boldsymbol{x}\notin\partial\mathcal{T}$, in spite of the fact that $1/r$ has an integrable singularity at $\boldsymbol{y}=\boldsymbol{x}$ when $\boldsymbol{x}\in\mathcal{T}$, as can be seen by excluding a small sphere of center $\boldsymbol{x}$ and radius ε from the integration over $\mathcal{T}$ and then passing to the limit as $\varepsilon \to 0$. This completes the proof.

The total magnetic energy due to the current $\boldsymbol{J}$ in the entire space

$$\mathcal{E}_m = \frac{1}{2}\mu_o \int_{\mathbb{R}^3} |\boldsymbol{H}(\boldsymbol{x})|^2 \boldsymbol{dx} \equiv \frac{1}{2}\mu_o \int_{\mathbb{R}^3} \boldsymbol{H}(\boldsymbol{x}) \cdot \boldsymbol{H}(\boldsymbol{x})\, \boldsymbol{dx} \tag{3.28}$$

can be expressed as a quadratic functional of $\boldsymbol{J}$. Indeed, since $\boldsymbol{H} = curl\ \boldsymbol{V}$, we have

$$\mathcal{E}_m = \frac{1}{2}\mu_o \int_{\mathbb{R}^3} \boldsymbol{H}(\boldsymbol{x}) \cdot curl\ \boldsymbol{V}(\boldsymbol{x})\, \boldsymbol{dx}$$

and applying the vector identity (1.48), with $curl\, \boldsymbol{H} = \boldsymbol{J}$, yields

$$\boldsymbol{H} \cdot curl\, \boldsymbol{V} \equiv \boldsymbol{V} \cdot curl\, \mathrm{H} + div(\boldsymbol{V} \wedge \boldsymbol{H}) = \boldsymbol{V} \cdot \boldsymbol{J} + div(\boldsymbol{V} \wedge \boldsymbol{H})$$

where, by force of eqs. (3.23) and (3.24),

$$\boldsymbol{V}(\boldsymbol{x}) \wedge \boldsymbol{H}(\boldsymbol{x}) = \mathbf{O}(|\boldsymbol{x}|^{-3}) \qquad \text{as } |\boldsymbol{x}| \to \infty$$

It follows that the term $div(\boldsymbol{V} \wedge \boldsymbol{H})$ gives no contribution to the integral in $\mathcal{E}_m$, and so

$$\mathcal{E}_m = \frac{1}{2}\mu_o \int_{\mathbb{R}^3} \boldsymbol{V}(\boldsymbol{x}) \cdot \boldsymbol{J}(\boldsymbol{x})\, \boldsymbol{dx} \tag{3.29}$$

Since $\boldsymbol{V}(\boldsymbol{x})$ is given by eq. (3.24), we finally obtain

$$\mathcal{E}_m = \frac{\mu_o}{8\pi} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\boldsymbol{J}(\boldsymbol{x}) \cdot \boldsymbol{J}(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} \mathbf{dxdy} \equiv \frac{\mu_o}{8\pi} \int_{\mathcal{T}} \int_{\mathcal{T}} \frac{\boldsymbol{J}(\boldsymbol{x}) \cdot \boldsymbol{J}(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} \mathbf{dxdy}$$

In this expression $\boldsymbol{J}$ is proportional to the Neumann vector field $\boldsymbol{N}$ which depends only on the geometry of $\mathcal{T}$, and therefore so does the normalized current density

$$\mathbb{J}(\boldsymbol{x}) := \frac{\boldsymbol{J}(\boldsymbol{x})}{I}$$

The magnetic energy $\mathcal{E}_m$ takes then the form

$$\mathcal{E}_m = \frac{1}{2} L I^2 \tag{3.30}$$

where the inductance of the conductor

$$L := \frac{\mu_o}{4\pi}\int_{\mathcal{T}}\int_{\mathcal{T}} \frac{\mathbb{J}(\boldsymbol{x})\cdot\mathbb{J}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\mathbf{dxdy} \tag{3.31}$$

is a positive quantity which depends only on the geometry of $\mathcal{T}$. By virtue of eqs. (3.31) and (3.29), the inductance can also be written in terms of the $L^2(\mathbb{R}^3)-$norm of $\boldsymbol{H}$

$$L = \frac{2\mathcal{E}_m}{I^2} \equiv \frac{\mu_o}{I^2}\|\boldsymbol{H}\|^2 \tag{3.32}$$

which is independent of I (since $\mathbf{H}$ is proportional to I).

If the conductor $\mathcal{T}$ is a thin wire all current lines are closed and are irreducible curves for $\mathcal{T}$. Let Γ denote one of these current lines lying on $\partial\mathcal{T}$, with tangent unit vector $\boldsymbol{t}(\boldsymbol{x})$, and let $S' \subset \mathcal{T}'$ denote an arbitrary surface with $\partial S' = \Gamma$. Then the magnetic flux linking Γ

$$\Phi := \int_{S'} \boldsymbol{B}\cdot\boldsymbol{n}\, dS \equiv \mu_o \int_{S'} \boldsymbol{H}\cdot\boldsymbol{n}_x dS_x$$

can be defined in an approximate sense as the magnetic flux linking $\mathcal{T}$.

Proposition 3.3.3 *If $\mathcal{T}$ is a thin wire then*

$$\Phi \cong LI\ ,\quad \mathcal{E}_m \cong \frac{1}{2}\Phi I \;\Leftrightarrow\; L \cong \frac{\Phi}{I} \tag{3.33}$$

Proof. If A is the average section area of $\mathcal{T}$ we have §1.1.3)

$$\mathbf{J} \cong \frac{I}{A}\mathbf{t} \qquad \Rightarrow \qquad A\mathbb{J}(\mathbf{x}) \cong \mathbf{t}(\mathbf{x}) \tag{3.34}$$

and by means of manipulations using eqs. (3.23), (3.24) and Stokes' theorem we find

$$\Phi \cong \mu_o \int_{S'} \boldsymbol{H}\cdot\boldsymbol{n}_x dS_x = \mu_o \int_{S'} curl_x \boldsymbol{V}\cdot\boldsymbol{n}_x dS_x = \mu_o \oint_\Gamma \boldsymbol{V}(\boldsymbol{x})\cdot\boldsymbol{t}(\boldsymbol{x})\, ds_x$$
$$\cong \mu_o\, A \oint_\Gamma ds_x \mathbb{J}(\boldsymbol{x})\cdot \frac{I}{4\pi}\int_{\mathcal{T}} \frac{\mathbb{J}(\boldsymbol{y})\boldsymbol{dy}}{|\boldsymbol{x}-\boldsymbol{y}|} \cong I\,\frac{\mu_o}{4\pi}\int_{\mathcal{T}}\int_{\mathcal{T}}\frac{\mathbb{J}(\boldsymbol{x})\cdot\mathbb{J}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\mathbf{dxdy} = LI$$

Eq. (3.31) yields then $\mathcal{E}_m = \Phi I$.

Since $\Phi \cong LI$ in the case of a thin wire, the Faraday induction law becomes

$$\frac{d\Phi}{dt} \cong -\frac{\mathcal{R}}{L}\Phi$$

and shows that the signs of Φ and $d\Phi/dt$ are opposite (Lenz's law). We also remark that the Biot-Savart formula has a counterpart in fluid mechanics, where **H** is replaced by the fluid velocity field $\boldsymbol{v}$, $\boldsymbol{J}$ is replaced by the vorticity *curl* $\boldsymbol{v}$, and the current lines by the vortex lines [8].

3.3.2 Ohm's law and energy balance.

We have already seen that for a toroidal conductor $\mathcal{T}$ in the presence of the impressed field $\boldsymbol{E}_{\text{imp}}$ with support $S = \Sigma_d$ the stationary energy balance equation for any domain $\Omega \subseteq \mathcal{T}$ is formally written as

$$\int_\Omega \gamma |\boldsymbol{E}|^2 \, dV + \int_{\partial\Omega} \boldsymbol{S} \cdot \boldsymbol{n} \, dS \; - \; P_{\text{imp}}[\Omega] \; = \; 0 \tag{3.35}$$

where $P_{\text{imp}}[\Omega]$ is given by eq. (3.16) with $S = \Sigma_d$ and is zero whenever $\Sigma_d \notin \Omega$.

By force of Proposition 3.3.1 and **H1**, the total current $I > 0$ is constant in $\mathcal{T} \backslash \Sigma_d$. We need a definition of resistance, impressed power, available power and power dissipated into heat for any domain $\Omega \subseteq \mathcal{T}$.

Definition 3.3.4 *The resistance of $\Omega \subseteq \mathcal{T}$ is defined by means of the $L^2(\Omega)$ norm of $\boldsymbol{J}$*

$$\mathcal{R}[\Omega] := \frac{1}{\gamma I^2} \big\| \boldsymbol{J} \big\|^2_{L^2(\Omega)} \; \equiv \; \frac{1}{\gamma I^2} \int_\Omega |\boldsymbol{J}(\boldsymbol{x})|^2 \, \boldsymbol{dx} \tag{3.36}$$

The *available power* is defined by the scalar product in $L^2(\Omega)$

$$\mathcal{P}_a[\Omega] := \int_\Omega \boldsymbol{E} \cdot \boldsymbol{J} \, dV \tag{3.37}$$

The *power dissipated into heat* by the Joule effect is $I^2\mathcal{R}[\Omega]$, and by (3.33)

$$I^2 \mathcal{R}[\Omega] = \gamma^{-1} \big\| \boldsymbol{J} \big\|^2_{L^2(\Omega)} \tag{3.38}$$

The *impressed power* (or power supplied by the generator) is $e_{\mathrm{i}}I$, where $e_{\mathrm{i}} > 0$ is the impressed e.m.f.

Note that the resistance $\mathcal{R}[\Omega]$ is an additive set function. For a thin wire of length l and section A the definition (3.3.4) yields the well-known value $\mathcal{R} = l/\gamma A$ (eq. (1.35)), the available power $\mathcal{P}_a$ equals VI, if V is the potential drop, and Ohm's law in the presence of an impressed e.m.f. is expressed by eq. (3.12).

We also define the modified Poynting vector

$$\tilde{\boldsymbol{S}} := \gamma^{-1}\boldsymbol{J} \wedge \boldsymbol{H} \tag{3.39}$$

which coincides with $\boldsymbol{S}{=}\boldsymbol{E}{\wedge}\boldsymbol{H}$ in any domain $\Omega \subseteq \mathcal{T} \backslash \Sigma_d$ which does not include the generator, so that $\boldsymbol{J}{=}\,\gamma\boldsymbol{E}$.

Proposition 3.3.5 *Let $\Omega \subseteq \mathcal{T}$ be any domain bounded by two equipotential sections Σ' and Σ'', distinct from Σ_d, having normals $\boldsymbol{n}{=}\boldsymbol{n}'$, $\boldsymbol{n}{=}\boldsymbol{n}''$ parallel to $\boldsymbol{J}$ and oriented so that $I > 0$. Define the potential drop due to $\boldsymbol{E}$*

$$V := \tilde{u}|_{\Sigma'} - \tilde{u}|_{\Sigma''}$$

(i) The power dissipated into heat $I^2\mathcal{R}[\Omega]$ satisfies

$$I^2\mathcal{R}[\Omega] = VI \ if \ \Sigma_d \notin \Omega \ , \ \ I^2\mathcal{R}[\Omega] = (V + e_i)I \ if \ \Sigma_d \in \Omega \tag{3.40}$$

so that Ohm's law reads

$$I\mathcal{R}[\Omega] = \begin{cases} V & if \ \Sigma_d \notin \Omega \\ V + e_i & if \ \Sigma_d \in \Omega \end{cases} \quad , \ e_i = I\mathcal{R}[\mathcal{T}] \tag{3.41}$$

(ii) The available power is equal to VI :

$$\mathcal{P}_a[\Omega] = VI \tag{3.42}$$

so that by combining eqs. (3.38) and (3.39) we have $\mathcal{P}_a[\Omega]{=}\ I^2\mathcal{R}[\Omega]$ if $\Sigma_d \notin \Omega$.

(iii) The power dissipated into heat is also equal to the opposite of the flux of the modified Poynting vector:

$$I^2\mathcal{R}[\Omega] = -\int_{\partial\Omega} \tilde{\boldsymbol{S}} \cdot \boldsymbol{n}\, dS \tag{3.43}$$

Hence in particular the total flux of $-\tilde{S}$ in $\mathcal{T}$ is equal to the impressed power:

$$-\int_{\partial\mathcal{T}} \tilde{\boldsymbol{S}} \cdot \boldsymbol{n}\, dS = e_i I \tag{3.44}$$

Proof. Since $\boldsymbol{J} = -\gamma\, grad\, u$ satisfies (3.17), we have by the definition of I

$$I^2\mathcal{R}[\Omega] = \frac{1}{\gamma}\int_\Omega |\boldsymbol{J}|^2 dV = -\int_\Omega \boldsymbol{J} \cdot grad\, u\, dV = -\int_\Omega div(\boldsymbol{J}u)\, dV = (u|_{\Sigma'} - u|_{\Sigma''})I$$

and by the definition of $\tilde{u}$

$$u|_{\Sigma'} - u|_{\Sigma''} = \begin{cases} \tilde{u}|_{\Sigma'} - \tilde{u}|_{\Sigma''} = V & if\ \Sigma_d \notin \Omega \\ \tilde{u}|_{\Sigma'} - (\tilde{u}|_{\Sigma''} - e_{\mathrm{i}}) = V + e_{\mathrm{i}} & if\ \Sigma_d \in \Omega \end{cases} \tag{3.45}$$

If $\Omega = \mathcal{T}$ then $\Sigma' = \Sigma''$, $V = 0$ and hence $e_{\mathrm{i}} = I\mathcal{R}[\mathcal{T}]$. This proves (i).

If $\Sigma_d \notin \Omega$ then $\mathbf{E} = \mathbf{J}/\gamma$ so that by force of (3.39) and (3.41) (cfr. also eq. (1.63))

$$\int_\Omega \boldsymbol{E} \cdot \boldsymbol{J}\, dV = \gamma^{-1} \int_\Omega |\boldsymbol{J}|^2 dV = I^2\mathcal{R}[\Omega] = VI$$

If $\Sigma_d \in \Omega$ then $\boldsymbol{E} = -grad\, u - \boldsymbol{E}_{\mathrm{imp}} = -grad\, u - e_{\mathrm{i}}\delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x})$ so that by (3.21) and (3.46)

$$\begin{aligned} \int_\Omega \mathbf{E} \cdot \boldsymbol{J}\, dV &= -\int_\Omega \boldsymbol{J} \cdot grad\, u\, dV - e_{\mathrm{i}} \int_\Omega \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{J}\, dV = \\ &= (V + e_{\mathrm{i}})I - e_{\mathrm{i}} \int_{\Sigma_d} \boldsymbol{n} \cdot \boldsymbol{J}\, dV = (V + e_{\mathrm{i}})I - e_{\mathrm{i}}I = VI \end{aligned} \tag{3.46}$$

as before. This proves (ii).

Since $\boldsymbol{J} = curl\, \boldsymbol{H} = -\gamma grad\, u$, we can rewrite the available power $\mathcal{P}_a[\Omega] = VI$ in the form (cfr. Exercise 15 of Chapter 1)

$$\mathcal{P}_a[\Omega] = \int_\Omega \boldsymbol{E} \cdot \boldsymbol{J} dV = -\int_\Omega curl\, \boldsymbol{H} \cdot grad\, u dV - e_{\mathrm{i}} \int_\Omega \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{J} dV \tag{3.47}$$

$$= -\frac{1}{\gamma}\int_{\Omega} div\,(\boldsymbol{J}\wedge\boldsymbol{H})\,dV - e_{\mathrm{i}}I = -e_{\mathrm{i}}I - \int_{\partial\Omega}\tilde{\boldsymbol{S}}\cdot\boldsymbol{n}\,dS$$

if $\Sigma_d \in \Omega$, and

$$\mathcal{P}_a[\Omega] = -\int_{\partial\Omega}\tilde{\boldsymbol{S}}\cdot\boldsymbol{n}\,dS$$

if $\Sigma_d \notin \Omega$. Since in both cases $\mathcal{P}_a[\Omega]= VI$, the assertion (iii) follows from eq. (3.41).

Remark 1. In the course of the proof of Proposition 3.3.5 we have seen that the impressed power is given by the integral

$$e_{\mathrm{i}}I = \int_{\Omega}\boldsymbol{E}_{\mathrm{imp}}\cdot\boldsymbol{J}\,dV \equiv -\gamma\int_{\Omega}\boldsymbol{E}_{\mathrm{imp}}\cdot grad\,u(\boldsymbol{x})\,dV \tag{3.48}$$

for any Ω with $\Sigma_d \in \Omega$, whereas from eqs. (3.16)

$$P_{\mathrm{imp}}[\Omega] := -\int_{\Omega}\boldsymbol{J}_{\mathrm{imp}}\cdot\boldsymbol{E}\,dV$$

where $\boldsymbol{J}_{\mathrm{imp}} = \gamma\boldsymbol{E}_{\mathrm{imp}}$ and $\boldsymbol{E}$ is given by eq. (3.18). It follows that

$$P_{\mathrm{imp}}[\Omega] = -e_{\mathrm{i}}I + \gamma\int_{\Omega}|\boldsymbol{E}_{\mathrm{imp}}|^2\,dV \tag{3.49}$$

However, since $|\boldsymbol{E}_{\mathrm{imp}}|$ is the Dirac distribution given by eq. (3.21), $|\boldsymbol{E}_{\mathrm{imp}}|^2$ is not defined, and therefore $P_{\mathrm{imp}}[\Omega]$ is not defined. Thus the energy balance equation (3.35) can be written only for domains Ω which do not include the generator Σ_d, so that $P_{\mathrm{imp}}[\Omega]= 0$. For such domains $\tilde{\boldsymbol{S}} = \boldsymbol{S}$ and (3.35) also follows from Theorem 3.3.5.

Remark 2. The relation $e_{\mathrm{i}} = I\mathcal{R}[\mathcal{T}]$in eq. (3.42 shows that e_{i} is uniquely determined by I, and vice-versa. Thus $u(\boldsymbol{x})$ and $\tilde{u}(\boldsymbol{x})$, that is $\boldsymbol{J}(\boldsymbol{x})$ and $\boldsymbol{E}(\boldsymbol{x})$, are uniquely determined for $\boldsymbol{x} \in \mathrm{T}\cup\partial\mathcal{T}$ if either the total current I or the impressed e.m.f. e_{i} is assigned. For instance, in the case of a homogeneous torus of conductivity γ we have (see the example in §3.1)

$$I = \gamma(a - \sqrt{a^2 - b^2})e_{\mathrm{i}}$$

and from the relation $e_{\mathrm{i}} = I\mathcal{R}[\mathcal{T}]$ we find that the resistance of the torus is

$$\mathcal{R}[\mathcal{T}] = \frac{1}{\gamma}\,\frac{1}{a - \sqrt{a^2 - b^2}}$$

Let $l = 2\pi a$ and $A = \pi b^2$ denote the length and cross section area of the torus, respectively. Then for $b \ll a$ $\mathcal{R}[\mathcal{T}]$ reduces to the usual formula for the resistance of a thin wire

$$\mathcal{R}[\mathcal{T}] \sim \frac{1}{\gamma}\frac{2\pi a}{\pi b^2} = \frac{l}{\gamma A}$$

which diverges in the infinitely thin wire limit $b \to 0$.

Remark 3. Eqs (3.32) and (3.36) together with eq (2.52) show that the resistance, inductance and capacity are proportional in their respective contexts to the $L^2(\mathbb{R}^3)-$norms of the normalized vector fields $\boldsymbol{J}/I, \boldsymbol{H}/I$ and $\boldsymbol{E}/V$, respectively:

$$\mathcal{R} = \frac{1}{\gamma I^2}\left\|\boldsymbol{J}\right\|^2 \quad , \quad L = \frac{\mu_o}{I^2}\left\|\boldsymbol{H}\right\|^2, \quad \mathbb{C} = \frac{\epsilon}{V^2}\left\|\boldsymbol{E}\right\|^2$$

3.3.3 The exterior electric field.

The issue whether or not a stationary resisting wire carrying a constant current I gives rise to a steady electric field $\boldsymbol{E}(\boldsymbol{x})$ outside it is still open to debate (see e.g. [1]). In the case of a toroidal conductor with a generator modeled by a potential jump and assigned current I the answer is affirmative, as already remarked by Heaviside.

Outside the conductor the electric field $\boldsymbol{E}(\boldsymbol{x})$ satisfies eqs. (3.7) and (3.8), so that

$$\boldsymbol{E}'(\boldsymbol{x}) = -grad\,\tilde{u}'(\boldsymbol{x}) \qquad (\mathbf{x} \in \mathcal{T}') \tag{3.50}$$

Proposition 3.3.6 *The potential $\tilde{u}'(\boldsymbol{x})$ is harmonic and one-valued in $\mathcal{T}'$.*

Proof. As $\boldsymbol{E}'(\boldsymbol{x})$ is solenoidal, $\Delta_3\tilde{u}'(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \mathcal{T}'$. Since $\partial\mathcal{T}$ is connected, by force of eq. (1.43) we have

$$\tilde{u}'_-(\boldsymbol{x}) = \tilde{u}(\boldsymbol{x}) \quad , \qquad \tilde{u}'_+(\boldsymbol{x}) = \tilde{u}(\boldsymbol{x}) + d \qquad \text{for } \mathbf{x} \in \partial\mathcal{T}\backslash\Sigma_d$$

where $d = [\tilde{u}']_{\partial\mathcal{T}}$ is the constant value of the discontinuity jump of $\tilde{u}'(\boldsymbol{x})$ across $\partial\mathcal{T}$ and $\tilde{u}(\mathbf{x})$ is the potential of the electric field $\boldsymbol{E}$ in $\mathcal{T}$. For any irreducible closed curve $\Gamma' \subset \partial\mathcal{T}\backslash\Sigma_d$ let S denote a surface contained in

$\mathcal{T}$ which does not intersect Σ_d and has boundary $\partial S = \Gamma'$. Since *curl* $\boldsymbol{E} = \boldsymbol{0}$ in $\mathcal{T}\backslash\Sigma_d$, and $\boldsymbol{E}\cdot\boldsymbol{t} = \boldsymbol{E}'\cdot\boldsymbol{t}$, by Stokes' theorem we have

$$\oint_{\Gamma'} \boldsymbol{E}' \cdot \boldsymbol{t}\, ds = \oint_{\Gamma'} \boldsymbol{E} \cdot \boldsymbol{t}\, ds = \int_S curl\, \boldsymbol{E} \cdot \boldsymbol{n}\, dS = 0$$

and so the period of $\tilde{u}'(\boldsymbol{x})$ is zero.

It follows that the potential $\tilde{u}'(\boldsymbol{x})$ is determined as the unique one-valued solution of the exterior Dirichlet problem

$$\begin{aligned} &\Delta_3 \tilde{u}'(\boldsymbol{x}) = 0 && (\boldsymbol{x} \in \mathcal{T}') \\ &\tilde{u}'(\boldsymbol{x}) = \tilde{u}(\boldsymbol{x}) + \delta && (\boldsymbol{x} \in \partial\mathcal{T}\backslash\Sigma_d) \\ &\tilde{u}'(\boldsymbol{x}) = O(\frac{1}{|\boldsymbol{x}|}) && |\mathbf{x}| \to +\infty \end{aligned} \tag{3.51}$$

with boundary data $\tilde{u}(\boldsymbol{x}) + \delta$ having a discontinuity jump equal to $\mathrm{e_i}$ when $\mathbf{x} \in \partial\mathcal{T}$ crosses the line $\partial\Sigma_d$. However, since by force of Proposition 3.1.1 $\tilde{u}(\boldsymbol{x})$ is bounded on $\partial\mathcal{T}$ and C^1 in $\partial\mathcal{T}\backslash\Sigma_d$, the solution $\boldsymbol{E}'(\boldsymbol{x}) = -grad\,\tilde{u}'(\boldsymbol{x})$ will be smooth in $\mathcal{T}'$ and continuous up to the boundary $\partial\mathcal{T}\backslash\Sigma_d$ if $\partial\mathcal{T}$ is smooth [2]. In order to determine it we must know δ, or in other words we must measure the value of $\tilde{u}'(\boldsymbol{x})$ at one (arbitrary) point of the boundary $\partial\mathcal{T}\backslash\Sigma_d$.

The solution $\tilde{u}'(\boldsymbol{x})$ of (3.52) can be decomposed into the sum

$$\tilde{u}'(\boldsymbol{x}) = U_o(\mathbf{x}) + \delta u_1(x)$$

of the solution $U_o(\boldsymbol{x})$ of the Dirichlet problem with $\delta = 0$ and of the capacitary potential $u_1(\boldsymbol{x})$ of $\partial\mathcal{T}$ (Chapter 2), so that

$$\boldsymbol{E}'(\boldsymbol{x}) = -grad\, U_o(\boldsymbol{x}) - \delta grad\, u_1(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \mathcal{T}')$$

According to Eq. (3.8), the surface charge density on the conductor's boundary is formally given by

$$\sigma = \epsilon \boldsymbol{E}_+ \cdot \boldsymbol{n} \equiv -\epsilon \frac{\partial U_o(\boldsymbol{x})}{\partial n} - \epsilon\delta \frac{\partial u_1(\boldsymbol{x})}{\partial n} \tag{3.52}$$

($\boldsymbol{x} \in \partial\mathcal{T}\backslash\Sigma_d$, $\boldsymbol{n} = \boldsymbol{n}(\boldsymbol{x})$) and if $\partial\mathcal{T}$ is smooth σ is well-defined and finite on $\partial\mathcal{T}\backslash\Sigma_d$. As $\tilde{u}$ is not constant in $\mathcal{T}$, $U_o(\boldsymbol{x})$ is not a capacitary potential for

$\partial\mathcal{T}$. It follows that $U_o(\boldsymbol{x}) + \delta u_1(\boldsymbol{x})$ cannot be constant in $\mathcal{T}'$ and its normal derivative cannot vanish identically on $\partial\mathcal{T}\backslash\Sigma_d$, even for particular values of δ.

Thus the electric field $\boldsymbol{E}'(\boldsymbol{x})$ is not identically zero outside the conductor and *the conductor boundary* $\partial\mathcal{T}$ is *electrically charged*. Thus $\boldsymbol{E}'(\boldsymbol{x})$ has a component orthogonal to the wire surface.

Remark 4. The Maxwell equation (1.25) with $\frac{\partial \boldsymbol{D}}{\partial t} \equiv \boldsymbol{0}$, applied to an irreducible curve $\Gamma' \subset \mathcal{T}'$ with unit tangent vector $\boldsymbol{t}'$ oriented according to the right-handed screw rule with respect to $\boldsymbol{n}$, yields

$$\oint_{\Gamma'} \boldsymbol{H} \cdot \boldsymbol{t}' \; ds = \int_S \boldsymbol{J} \cdot \boldsymbol{n} \; dS \qquad\qquad (\partial S = \Gamma') \tag{3.53}$$

where $S \subset \mathcal{T}'$ is any surface bounded by Γ' and having normal $\boldsymbol{n}$ such that $\boldsymbol{J}\cdot\boldsymbol{n} > 0$. Since S necessarily cuts $\mathcal{T}$, the right-hand side of eq. (3.53) is equal to $I > 0$. Hence the Biot-Savart magnetic field (3.23) can be written for $\boldsymbol{x} \in \mathcal{T}'$ as the gradient

$$\boldsymbol{H}(\boldsymbol{x}) = -grad\, v(\boldsymbol{x})$$

of a potential $v(\boldsymbol{x})$ which is harmonic and many-valued in $\mathcal{T}'$, with period

$$p_b := -\oint_{\Gamma'} \boldsymbol{H} \cdot \boldsymbol{t}\, ds = -I$$

Moreover, v satisfies the asymptotic condition at infinity

$$v(\boldsymbol{x}) = O(\,|\boldsymbol{x}|^{-1}) \qquad \text{as } |\boldsymbol{x}| \to \infty \tag{3.54}$$

(uniformly with respect to direction) as a consequence of the condition (3.5) for $\boldsymbol{H}$. If, on the other hand, $\boldsymbol{t}'$ is oriented according to the left-handed screw rule with respect to $\boldsymbol{n}$, the sign in the Maxwell equation (3.54) must be changed and we obtain

$$p_b := -\oint_{\Gamma'} \boldsymbol{H} \cdot \boldsymbol{t}\, ds = I \tag{3.55}$$

This remark will turn out to be useful when dealing with infinitely thin wires.

3.4 The inductance matrix

We consider now N homogeneous non-magnetic toroidal conductors $\mathcal{T}_j$, with conductivities γ_j $(0 < \gamma_j < +\infty)$, carrying electric currents of density $\boldsymbol{J}_j(\boldsymbol{x})$ $(j = 1, ..., N)$, and surrounded by an unbounded homogeneous uncharged dielectric of permittivity ϵ. As before, we have

$$\mu(\boldsymbol{x}) = \mu_o \ , \quad \rho(\boldsymbol{x}) \equiv 0 \ \ \forall\, \boldsymbol{x} \in \mathbb{R}^3$$

We assume that the conductors are disjoint and unknotted, and we consider for simplicity the case $N = 2$; the extension to $N > 2$ is immediate. The currents $\boldsymbol{J}_k(\boldsymbol{x})$ satisfy eq. (3.1) in $\mathcal{T}_k$, and therefore are given by the corresponding interior Neumann vector fields

$$\boldsymbol{J}_k(\boldsymbol{x}) = \mathbb{N}_k(\boldsymbol{x}) \equiv -\gamma grad\, u_k \qquad \text{for } \boldsymbol{x} \in \mathcal{T}_k \cup \partial\mathcal{T}_k$$

where u_k is the $k-$th interior potential of $\boldsymbol{J}_k/\gamma_k$ in $\mathcal{T}_k$ $(k = 1, 2)$. The total current in $\mathcal{T}_k$ is given by the integral

$$I_k = \int_{\Sigma_k} \boldsymbol{n} \cdot \boldsymbol{J}_k \, dS$$

for a generic section Σ_k of $\mathcal{T}_k$ and, if the potentials $u_k(\boldsymbol{x})$ satisfy assumption H1, we can assume that $I_k > 0$ $(k = 1, 2)$. Letting

$$\boldsymbol{J}(\boldsymbol{x}) := \begin{cases} \boldsymbol{J}_k(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \mathcal{T}_k \\ \boldsymbol{0} & \text{otherwise} \end{cases}$$

the magnetic field $\boldsymbol{H}$ satisfies formally the same equations (3.22), and is therefore given by the extended Biot-Savart law

$$\boldsymbol{H}(\boldsymbol{x}) = curl\ \boldsymbol{V}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^3$$

with vector potential defined by

$$\text{(3.56)} \qquad \boldsymbol{V}(\boldsymbol{x}) := \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{\boldsymbol{J}(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} \, \boldsymbol{dy} \equiv \sum_{k=1}^{2} \frac{1}{4\pi} \int_{\mathcal{T}_k} \frac{\boldsymbol{J}_k(\boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|} \, \boldsymbol{dy}$$

The magnetic energy can still be expressed in the form (3.29), so that here

$$\mathcal{E}_m = \frac{1}{2}\mu_o \int_{\mathbb{R}^3} \boldsymbol{V}(\boldsymbol{x}) \cdot \boldsymbol{J}(\boldsymbol{x}) \, \boldsymbol{dx} \ \equiv \ \frac{1}{2}\mu_o \sum_{j=1}^{2} \int_{\mathcal{T}_j} \boldsymbol{V}(\boldsymbol{x}) \cdot \boldsymbol{J}_j(\boldsymbol{x}) \, \boldsymbol{dx}$$

Substituting (3.56) into the previous equation we find

$$\mathcal{E}_m = \frac{\mu_o}{8\pi}\sum_{j,k=1}^{2}\int_{\mathcal{T}_j}\int_{\mathcal{T}_k}\frac{\mathbb{J}_k(\boldsymbol{y})\cdot\mathbb{J}_j(\boldsymbol{x})}{|\boldsymbol{x}-\boldsymbol{y}|}\boldsymbol{dx}\,\boldsymbol{dy} \;\equiv\; \frac{1}{2}\sum_{j,k=1}^{2}L_{jk}I_jI_k \tag{3.57}$$

where $\mathbb{J}_k$ is the $k-$th normalized current density

$$\mathbb{J}_k(\boldsymbol{x}) \;:=\frac{\boldsymbol{J}_k}{I_k}$$

$(k = 1, 2)$, and

$$L_{jk} := \frac{\mu_o}{4\pi}\int_{\mathcal{T}_j}\int_{\mathcal{T}_k}\frac{\mathbb{J}_k(\boldsymbol{y})\cdot\mathbb{J}_j(\boldsymbol{x})}{|\boldsymbol{x}-\boldsymbol{y}|}\mathbf{dx}\,\mathbf{dy} \tag{3.58}$$

$(j, k = 1, 2)$ are the entries of the inductance matrix. This matrix depends only on the geometry of the conductors. L_{kk} is the self-inductance of $\mathcal{T}_k$ $(k = 1, 2)$ and L_{12} is called the mutual inductance. Clearly $L_{kk} > 0$, $L_{12} = L_{21}$ and, since $\mathcal{E}_m > 0$ for $I_1I_2 > 0,$ the inductance matrix is symmetric and positive definite. Therefore the mutual inductance satisfies

$$L_{12} = L_{21} < \sqrt{L_{11}L_{22}} \tag{3.59}$$

If the conductors $\mathcal{T}_k$ are thin wires, the magnetic flux linking $\mathcal{T}_k$ is a linear combination of the currents with approximate coefficients L_{kj}

$$\Phi_k := \mu_o\int_{S_k}\boldsymbol{H}\cdot\boldsymbol{n}_x dS_x \;\cong\sum_{j=1}^{2}L_{kj}I_j \qquad\qquad (k = 1, 2) \tag{3.60}$$

(see Prpoposition 3.3.3) and the magnetic energy can be approximated by

$$\mathcal{E}_m \cong \frac{1}{2}\sum_{k=1}^{2}\Phi_kI_k$$

3.5 Magnetic field due to an infinitely thin wire.

We will now show that the Biot-Savart law (1.18) for an infinitely thin wire Γ follows from the Maxwell equations. This can be done by a limit process, using the expression (1.11) for the linear current density $\boldsymbol{J}$

$$\boldsymbol{J}(\boldsymbol{x}) = I\delta_\Gamma(\boldsymbol{x})\boldsymbol{t}(\boldsymbol{x}) \tag{3.61}$$

where $\boldsymbol{t}(\boldsymbol{x})$ is the tangent vector to Γ at the point $\boldsymbol{x}$. Substituting this expression in eq. (3.23) we obtain ($\boldsymbol{x} \in \mathbb{R}^3 \backslash \Gamma$)

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{I}{4\pi}\int_{\mathcal{T}} \delta_\Gamma(\boldsymbol{y})\boldsymbol{t}(\boldsymbol{y}) \wedge \frac{\boldsymbol{x}-\boldsymbol{y}}{|\boldsymbol{x}-\boldsymbol{y}|^3}\, \boldsymbol{dy} \equiv \frac{I}{4\pi}\int_\Gamma \boldsymbol{t}(\boldsymbol{y}) \wedge \frac{\boldsymbol{x}-\boldsymbol{y}}{|\boldsymbol{x}-\boldsymbol{y}|^3}\, ds_{\mathbf{y}} \tag{3.62}$$

which is the integrated form of the Biot-Savart law (1.18). In this way $\boldsymbol{H}(\boldsymbol{x})$ is expressed as the curl of the vector potential

$$\boldsymbol{V}(\boldsymbol{x}) = \frac{I}{4\pi}\oint_\Gamma \frac{\boldsymbol{t}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\, ds_y$$

An alternative derivation, which will be useful in the case of a solenoid (§3.6), is based on the fact that for an infinitely thin wire there exists a many-valued scalar magnetic potential $v(\boldsymbol{x})$ in the space surrounding the wire, as anticipated in Remark 4. The magnetic field can then be represented in the form of a gradient

$$\boldsymbol{H}(\boldsymbol{x}) = -grad\, v(\boldsymbol{x}) \qquad (\boldsymbol{x} \in \mathbb{R}^3 \backslash \Gamma) \tag{3.63}$$

where the potential $v(\boldsymbol{x})$ is harmonic and many-valued in $\mathbb{R}^3 \backslash \Gamma$, with period

$$p_b := -\oint_{\Gamma_e} \boldsymbol{H} \cdot \boldsymbol{t}'\, ds = I$$

Here Γ_e is any closed curve linking Γ with unit tangent $\mathbf{t}'$ oriented according to the left-hand screw rule with respect to $\mathbf{t}$ (see eq. (3.56)). Thus, the many-valued potential v has an infinite number of branches: by making k turns around Γ, v resumes its initial value augmented or diminished by $kp_b = kI$, $k = 1, 2, \ldots$. The magnetic field itself is regular and one-valued, since the difference between any two branches of v is a constant. We may then introduce an arbitrary branch surface Σ_b bounded by Γ, so that the domain $\mathbb{R}^3 \backslash \Sigma_b$ becomes simply connected and the potential v is restricted to a one-valued branch in $\mathbb{R}^3 \backslash \Sigma_b$ which has a constant discontinuity jump

$$v_+ - v_- \equiv [v]_{\Sigma_b} = I \tag{3.64}$$

across the two sides $\pm$ of the branch surface Σ_b . Here v_- and v_+ denote the values of v on Σ_b at the initial and final points of Γ_e, respectively. If the normal $\mathbf{n}$ to Σ_b is oriented from the $-$ side to the $+$ side, as usual, it turns out that this orientation also satisfies the right-hand screw rule with respect to $I\mathbf{t}$ (see Fig. 3.2). By introducing the branch surface Σ_b, the determination

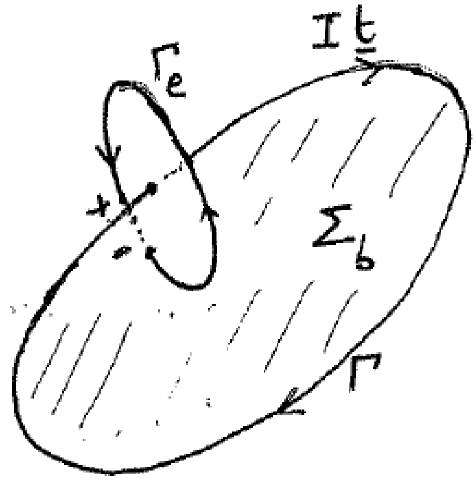

Figure 3.2: Branch surface for an infinitely thin wire

of the potential $v(\boldsymbol{x})$ reduces to finding the solution of the boundary value problem (summation problem)

$$\begin{aligned} &\Delta_3 v = 0 && \text{in } \mathbb{R}^3\backslash\overline{\Sigma_b} \\ &[v]_{\Sigma_b} = I && \text{across } \Sigma_b \\ &v = O(|\boldsymbol{x}|^{-1}) && \text{as } |\boldsymbol{x}| \to +\infty \ \text{(uniformly)} \end{aligned} \tag{3.65}$$

where the condition at infinity corresponds to the asymptotic condition for $\boldsymbol{H}$ assumed in Proposition 3.3.2

Proposition 3.5.1 *For any fixed Σ_b, with $\partial\Sigma_b = \Gamma$, the solution $v \in C^2(\mathbb{R}^3\backslash\Gamma)$ of (3.65) is unique and is given by the double layer potential $\mathcal{W}_I(\boldsymbol{x})$ with constant density I*

$$v(\boldsymbol{x}) = \frac{I}{4\pi}\int_{\Sigma_b} \frac{\partial}{\partial n_y}(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|})\, dS_y\,, \qquad \boldsymbol{x} \in \mathbb{R}^3\backslash\overline{\Sigma}_b \tag{3.66}$$

Proof. If $I = 0$ the potential v is one-valued and the branch surface disappears, so that $v \equiv 0$ is the sole solution of (3.65) by virtue of Liouville's theorem for harmonic functions [2]. This proves uniqueness. The fact that the solution is represented by (3.66) follows from the properties of the double layer potential stated in §2.1.

Proposition 3.5.2 *The magnetic field*

$$\boldsymbol{H}(\boldsymbol{x}) = -\,grad\,\mathcal{W}_I(\boldsymbol{x}) \;\equiv -\frac{I}{4\pi} grad_x \int_{\Sigma_b} \frac{\partial}{\partial n_y}\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) dS_y \tag{3.67}$$

is independent of the choice of the branch surface Σ_b *and can be written in the Biot-Savart form (3.62) for* $\boldsymbol{x}\in \mathbb{R}^3\backslash\Gamma$. $\boldsymbol{H}(\boldsymbol{x})$ *is smooth for* $\boldsymbol{x}\in \mathbb{R}^3\backslash\Gamma$, *including the branch surface* Σ_b, *but is singular as* $\boldsymbol{x}\rightarrow \Gamma$. *In particular*

$$|\boldsymbol{H}(\boldsymbol{x})| = O\left(\frac{1}{\varrho}\right) \qquad as\ \varrho \rightarrow 0$$

if ϱ *is the distance of the point* $\boldsymbol{x}$ *from the wire.*

Proof. For any open surface S, the identity

$$grad_x \int_S \frac{\partial}{\partial n_y}\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) dS_y \equiv -curl_x \oint_{\partial S} \frac{\boldsymbol{t}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|}\, ds_y \tag{3.68}$$

holds for any $\boldsymbol{x}\in \mathbb{R}^3\backslash\overline{S}$. Indeed, from Stokes's theorem (ST2) of Chapter 1 applied to the function $f = 1/r$, $r = |\boldsymbol{x}-\boldsymbol{y}|$:

$$\int_S \boldsymbol{n}(\boldsymbol{y}) \wedge grad\frac{1}{r}\, dS_y \;=\; \oint_{\partial S} \boldsymbol{t}(\boldsymbol{y})\,\frac{1}{r}\, ds_y$$

and from the vector identity [3]

$$curl_x\,(\boldsymbol{n}(\boldsymbol{y}) \wedge \boldsymbol{w}(\boldsymbol{x})) \equiv \boldsymbol{n}(\boldsymbol{y})\, div_x \boldsymbol{w}(\boldsymbol{x}) \,-\, (\boldsymbol{n}(\boldsymbol{y}) \cdot grad_x)\boldsymbol{w}(\boldsymbol{x})$$

with $\boldsymbol{w}= grad_x(1/r) \equiv -grad_y(1/r)$, we obtain

$$curl_x \oint_{\partial S} \boldsymbol{t}(\boldsymbol{y})\frac{1}{r}\, ds_y = curl_x \int_S \boldsymbol{n}(\boldsymbol{y}) \wedge grad_y\frac{1}{r}\, dS_y \;= -\int_S curl_x\left(\boldsymbol{n}(\boldsymbol{y}) \wedge grad_x\frac{1}{r}\right) dS_y$$

$$= -\oint_S \boldsymbol{n}(\boldsymbol{y})\Delta_3\frac{1}{r}\, dS_y + \oint_S (\boldsymbol{n}(\boldsymbol{y}) \cdot grad_x) grad_x\frac{1}{r}\, dS_y = -grad_x \oint_S \frac{\partial}{\partial n_y}\frac{1}{r}\, dS_y$$

[3] in general $curl\,(\mathbf{a}\wedge\mathbf{b}) \equiv \mathbf{a}\, div\mathbf{b} - \mathbf{b}\, div\mathbf{a} + (\mathbf{b}\cdot grad\,)\mathbf{a} - (\mathbf{a}\cdot grad\,)\mathbf{b}$

where we have used the fact that $1/r$ is harmonic for $r \neq 0$. Applying this identity to eq. (3.67), where $S = \Sigma_b$ with $\partial\Sigma_b = \Gamma$, yields the Biot-Savart formula (3.62)

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{I}{4\pi} curl_x \oint_{\partial\Sigma_b} \frac{\boldsymbol{t}(\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|} ds_y \equiv \frac{I}{4\pi} \int_\Gamma \boldsymbol{t}(\boldsymbol{y}) \wedge \frac{\boldsymbol{x}-\boldsymbol{y}}{|\boldsymbol{x}-\boldsymbol{y}|^3} ds_y \tag{3.69}$$

Hence $\boldsymbol{H}$ is independent of the choice of the branch surface Σ_b, and eq. (3.69) holds for all $\boldsymbol{x} \in \mathbb{R}^3 \backslash \Gamma$.

The branch surface Σ_b can be interpreted as a virtual magnetic double layer with constant dipole density I, and the identity (3.68) embodies the

Ampère's equivalence principle: the magnetic field $\boldsymbol{H}$ generated by an infinitely thin wire Γ carrying a steady current I is the same as that due to a magnetic double layer with constant dipole intensity I distributed over any surface Σ bounded by Γ.

If the area $\mathcal{A} = d\Sigma$ tends to zero (or equivalently if $|\boldsymbol{x}| \to \infty$) from eq. (3.66) we obtain the limit expression [4]

$$v(\boldsymbol{x}) \sim \frac{I\mathcal{A}}{4\pi} \frac{\partial}{\partial n_y} \left(\frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\right) = \frac{I}{4\pi} \frac{\boldsymbol{n}_y \cdot \boldsymbol{r}}{r^3} d\Sigma \qquad (\mathbf{r} = \boldsymbol{x} - \boldsymbol{y})$$

which coincides with the potential at the point $\boldsymbol{x}$ of a magnetic dipole

$$v(\boldsymbol{x}) = \frac{\boldsymbol{m} \cdot \boldsymbol{r}}{4\pi r^3}$$

situated at the point $\boldsymbol{y}$ with moment $\boldsymbol{m} = I\mathcal{A}\boldsymbol{n} \equiv I\boldsymbol{n}_y d\Sigma$ (cfr. eq. (1.15) with $\mathbf{b} = \boldsymbol{n}$). Thus every magnetic dipole is equivalent to a small (infinitesimal) current loop. If one thinks of a branch surface Σ_b as being "tiled" by such elements $d\Sigma$, the interior currents cancel out leaving only the contribution (3.69) of the current I on the boundary $\Gamma = \partial\Sigma$. This is the intuitive content of Proposition 3.5.2 and of eq. (3.68).

Example 1. (the circular loop). The Ampère equivalence principle enables us to derive easily a formula for the magnetic field due to a circular loop Γ of wire carrying a current I. The magnetic potential is given by eq. (3.66) and is proportional to the solid angle (with sign) subtended by Σ_b at the point $\boldsymbol{x}$ (see §2.1.2). Suppose that the loop is centered at the point

[4] we recall that $f \sim g$ means that f/g tends to 1 (see §1.1)

$(0,0,z)$ with binormal $\mathbf{b}{=}\mathbf{c}_3$ and that Σ_b is taken as the circle bounded by Γ in the $z-$plane. Then if $\boldsymbol{x}{=}\,(0,0,Z)$ the integral in (3.66) is easily computed and yields (Exercise 4)

$$v(0,0,Z)=\frac{I}{2}\Big[\frac{Z-z}{|Z-z|}-\frac{z}{\sqrt{R^2+(Z-z)^2}}\Big]$$

so that the magnetic field on the axis of the loop is given by

$$\mathbf{H}(0,0,Z)=\frac{I\ R^2\ \mathrm{c}_3}{2(R^2+(Z-z)^2)^{3/2}} \tag{3.70}$$

Example 2. The Biot-Savart magnetic field for an infinite straight wire

$$\boldsymbol{H}(\boldsymbol{x})=\frac{I}{2\pi\varrho}\boldsymbol{\tau}$$

can be written in the form (3.67), with branch surface Σ_b given by any half-plane bounded by the wire (see and Exercise 10 of Chapter 2).

Remark 5. Ampère's principle (3.68) can be generalized to double layers of variable density $\nu(\boldsymbol{y})$:

$$curl_x\int_S \nu(\boldsymbol{y})\mathbf{n}(\boldsymbol{y})\wedge grad\frac{1}{r}dS_{\mathbf{y}}+grad_{\mathbf{x}}\int_S \nu(\boldsymbol{y})\frac{\partial}{\partial n_y}\frac{1}{r}dS_{\mathbf{y}}\equiv -curl_{\mathbf{x}}\oint_{\partial S}\frac{\boldsymbol{t}(\boldsymbol{y})\nu(y)}{r}\,ds_{\mathbf{y}}$$

$(r=|\boldsymbol{x}-\boldsymbol{y}|)$. In fluid mechanics this identity is interpreted as the equivalence of doublets and vortex layers [2,8].

The following propositions clarify the detailed singular behavior of the Biot-Savart field on Γ, that is of the gradient of a double layer with constant density distributed on a surface with boundary Γ (see also Theorem 2.1.5 and Exercise 10 of Chapter 2).

Proposition 3.5.3 *Denote by* $(\boldsymbol{n},\boldsymbol{b},\boldsymbol{t})$ *the Frenet trihedron at a point* $\boldsymbol{x}_o\in\Gamma$, $\kappa=\kappa(\boldsymbol{x}_o)$ *the curvature,*

$$\boldsymbol{\tau}=-sin\varphi\,\boldsymbol{n}+cos\varphi\,\boldsymbol{b}$$

the transverse unit vector in the normal plane $(\boldsymbol{n},\boldsymbol{b})$ *at* $\boldsymbol{x}_o$ *[19]. Choose a cartesian reference frame with origin at* $\boldsymbol{x}_o$ *and axes* $(\boldsymbol{n},\boldsymbol{b},\boldsymbol{t})$, *so that* $\boldsymbol{c}_1=\boldsymbol{n}$ *is* *the principal normal,* $\boldsymbol{c}_2 = \boldsymbol{b}$ *is the binormal,* $\boldsymbol{c}_3 = \boldsymbol{t}$ *the unit tangent vector to* Γ *at* $\boldsymbol{x}_o$*. Consider a point* $\boldsymbol{x} = (x_1, x_2, 0) = (\varrho cos\varphi,\ \varrho sin\varphi,\ 0)$ *and let* ϱ *tend to zero, so that* $\boldsymbol{x}$ *approaches* $\boldsymbol{x}_o$ *in the normal plane. Then the Biot-Savart field (3.67) or (3.69) satisfies*

$$\boldsymbol{H}(\varrho\, cos\varphi,\ \varrho sin\varphi,\ 0) = \frac{I}{2\pi\varrho}\,\boldsymbol{\tau} - \frac{I\kappa}{4\pi} log(\varrho)\,\boldsymbol{b}\ + O(1) \quad as\ \ \varrho \to 0 \tag{3.71}$$

For the proof, see [8], p. 510.

Proposition 3.5.4 *The Biot-Savart field* $\boldsymbol{H} \notin L^2_{loc}(\mathbb{R}^3 \backslash \Gamma)$.

Proof. The singular behavior (3.71) implies that $\boldsymbol{H}$ is not square summable in a neighborhood of Γ. In short, $\boldsymbol{J}(\boldsymbol{x}) = I\delta_\Gamma(\boldsymbol{x})\boldsymbol{t}(\boldsymbol{x})$ belongs to the Sobolev space $H^{-1-\varepsilon}(\mathbb{R}^3)$ for $\varepsilon > 0$ [13], hence $\boldsymbol{H} = curl^{-1}\boldsymbol{J} \in H^{-\varepsilon}(\mathbb{R}^3)$ is not in $L^2(\mathbb{R}^3)$.

These results show that the model of linear currents in infinitely thin wires, although useful for discussing certain properties of the magnetic field of a toroidal conductor of very small cross section (see e.g. §3.6), is intrinsically inconsistent [5]. For example, the local magnetic energy

$$\mathcal{E}_m(K) := \frac{1}{2}\mu_o \int_K |\boldsymbol{H}(\boldsymbol{x})|^2 \boldsymbol{dx}$$

is infinite in any neighborhood K of the wire Γ, a physical nonsense. The (self-) inductance of the wire should be defined by the integral, indipendent of I and of the choice of Σ_b

$$L := \mu_o \int_{\mathbb{R}^3} |grad\, \mathcal{W}_1(\boldsymbol{x})|^2\, dV$$

and should satisfy the relations (cf. eqs. (3.30))

$$\Phi = LI \qquad , \qquad \mathcal{E}_m = \frac{1}{2}LI^2 \equiv \frac{1}{2}\Phi I \tag{3.72}$$

where

$$\Phi := \mu_o \int_{\Sigma_b} \boldsymbol{H} \cdot \boldsymbol{n}\, dS \tag{3.73}$$

[5] the same inconsistencies arise for the model of the line vortex in hydrodynamics ([8], pp. 509-511)

is the magnetic flux linking Γ, and

$$\mathcal{E}_m := \frac{1}{2}\mu_o \int_{\mathbb{R}^3} |\boldsymbol{H}(\boldsymbol{x})|^2 \boldsymbol{dx}$$

is the total magnetic energy. Unfortunately all these integrals are divergent, as a consequence of Proposition 3.5.3 and Proposition 3.5.4 (see Exercise 9). Similarly, the integrals (3.31) and (3.58) for L_{kk} diverge in the limit of an infinitely thin wire (Exercise 5). Only the integral (3.58) for the mutual inductance $L_{jk}\,(j \neq k)$ remains finite, as the integrand is obviously regular.

A further inconsistency arises from the evaluation of the self-induced action, that is the resultant force exerted by the Biot-Savart field on the wire itself. We have seen in Proposition 3.5.3 that, if $\kappa(\boldsymbol{x}_o) \neq 0, \boldsymbol{H(x)}$ has, besides the leading singularity $I\boldsymbol{\tau}/2\pi\varrho$, an additional logarithmic singularity as $\boldsymbol{x}\to\boldsymbol{x}_o$, directed along the binormal $\mathbf{b}(\boldsymbol{x}_o)$. This logarithmic term yields in general a non-vanishing self-induced force [6]. Consider eq. (1.13)

$$d\boldsymbol{F} = \mu_o \boldsymbol{J} \wedge \boldsymbol{H}\; dV \cong \mu_o I \boldsymbol{t} \wedge \boldsymbol{H}\, ds \tag{3.74}$$

written for a current tube $dV \cong \mathcal{A}ds$ centered on the wire Γ, where $\mathcal{A}$ is the area of the tube cross-section, supposed circular. The self-induced force for a wire Γ can be computed by substituting (3.71) in (3.74), integrating over the current tube and then passing to the limit as $\mathcal{A} \to 0$:

$$\boldsymbol{F} = \frac{\mu_o I^2}{2\pi} \lim_{A\to 0}\left[\int_{\Gamma\times\mathcal{A}} \boldsymbol{t}\wedge\boldsymbol{\tau}\,\frac{ds\,d\mathcal{A}}{\varrho} + \frac{1}{2}\int_{\Gamma\times\mathcal{A}} \kappa\, log(\varrho)\,\mathbf{b}\wedge\boldsymbol{t}\, ds\, d\mathcal{A}\right]$$

Since $\mathbf{b}\wedge\boldsymbol{t}{=}\boldsymbol{n}$ and $\boldsymbol{t}\wedge\boldsymbol{\tau}$ is the radial unit vector in the plane orthogonal to Γ, for reasons of symmetry the first integral is always zero while the second is zero if $\kappa = 0$ (rectilinear wire) or κ =constant (circular loop). In all other cases the self-induced action $\boldsymbol{F}$ does not vanish, a physical inconsistency which is contrary to what happens in the case of a toroidal conductor (Exercise 6).

In order to give a meaning to eqs. (3.72) one must suppose that the section of the wire is small but finite, and use the formulas established in

[6] in hydrodynamics this logarithmic term implies that a curvilinear line vortex moves with infinite speed along the binormal, a rather unbecoming behavior ([8], p.511)

section 3.3 for toroidal conductors. In this way eqs. (3.72) can be rigorously established on the basis of (3.33) and can be further extended to quasi-stationary fields (see §3.7).

Remark 6. A similar inconsistency arises, as discussed in §§1.1.3 and 1.3.2 , when considering the electric field. According to eqs. (3.18) and (3.21), the electric field inside a thin wire of small but finite cross section area $\mathcal{A}$, corresponding to the linear current density (3.61), $\boldsymbol{J}(\boldsymbol{x}) = I\mathcal{A}^{-1}\,\boldsymbol{t}(\boldsymbol{x})$, has the form

$$\boldsymbol{E}\,(\boldsymbol{x}(s)) = (\gamma\mathcal{A})^{-1}\,I\,\boldsymbol{t}(\boldsymbol{x}(s)) - e_{\mathrm{i}}\delta(s-s_o)\,\boldsymbol{t}(\boldsymbol{x}_o)$$

where s is arc-length, e_{i} is the impressed e.m.f. and $e_{\mathrm{i}}\delta(s-s_o)\,\boldsymbol{t}(\boldsymbol{x}_o)$ is the impressed electric field due to a generator at the point $\mathbf{x}_o = \boldsymbol{x}(s_o)$ of the circuit. Clearly, this expression admits no finite limit as $\mathcal{A} \to 0$. Hence the issue whether or not an electric field $\boldsymbol{E}$ exists outside (or inside!) an infinitely thin wire carrying a constant current is meaningless. For a wire of small but finite cross section the answer is affirmative, as we have shown above.

3.6 Magnetic field of a solenoid

We apply the model of linear currents to the evaluation of the magnetic field of a solenoid, consisting of a large number N of helical turns of an infinitely thin wire carrying a steady current I and closely wound around an insulating non-magnetic core in the shape of a circular cylinder with directrices parallel, say, to the unit vector $\mathbf{c}_3$ of the $z-$axis, oriented according to the right-hand screw rule with respect to the positive flow of current. We assume the same geometry as in Fig. 2.3 for the bar magnet. In particular, R is the cylinder radius, $2l$ the solenoid height, $z = \pm l$ the height of the two bases. If

$$N_o = dN/dz = N/2l$$

denotes the number of turns per unit length, we assume $N_o l \gg 1$ and $l \gg R$. As before, we can suppose that $\mu = \mu_o$ everywhere.

The branch surface Σ_b is now a helical surface of N sheets Σ_{bi} ($i = 1, ..., N$). Eq. (3.66) and the principle of superposition imply that the magnetic potential can be written as the sum

$$v = \sum_{i=1}^{N} v_i(\boldsymbol{x})$$

of N double layers potentials v_i distributed on the $i-$th sheet Σ_{bi} , respectively:

$$v_i(\boldsymbol{x}) := \frac{I}{4\pi}\int_{\Sigma_{bi}} \frac{\partial}{\partial n_y}(\frac{1}{r})\; dS_y \qquad\qquad (i = 1, ..., N)$$

An exact calculation of the magnetic intensity field $\boldsymbol{H}= -grad\, v$ using these formulae is cumbersome, especially in the interior of the solenoid. An approximate picture can be obtained if we remark that, under the assumption $N_o l \gg 1,$ the solenoid can be replaced by N_o circumferential loops per unit length and Σ_{bi} by the circular cross section Σ_i at the height $z = z_i$, oriented by taking $\mathbf{n}$=$\mathbf{c}_3$. The potentials v_i can then be written as

$$v_i(\boldsymbol{x}) := \frac{I}{4\pi}\int_{\Sigma_i} \frac{\partial}{\partial n_y}(\frac{1}{r})\; dS_y \tag{3.75}$$

By force of Propositions 3.5.1, and 3.5.2 and of eq. (3.64), each v_i satisfies the jump relations

$$[v_i]_{S_j} = I\delta_{ij} \qquad (i, j = 1, ..., N)$$

where δ_{ij} denotes the Kronecker delta, and so $v = \sum_{i=1}^{N} v_i(\boldsymbol{x})$ satisfies

$$[v]_{\Sigma_j} = I \qquad\qquad (j = 1, ..., N)$$

as required. These relations imply that

$$v(z + dz) - v(z) = I\, dN \quad , \quad -l \leq z < l$$

where the discrete quantity dN assumes integer values equal to the number of sections Σ_i included between z and $z + dz$. Since $N_o l \gg 1$, the integer dN can be approximately replaced with the continuous quantity $N_o dz$, which vanishes as dz tends to zero, Σ_i by $\Sigma(z)$, and $\partial/\partial n_y$ by $\partial/\partial z$. The sum of the double layers (3.75)

$$v(\boldsymbol{x}) = \frac{I}{4\pi}\sum_{i=1}^{N} dN_i \int_{\Sigma_i} \frac{\partial}{\partial n_y}\frac{1}{r}\; dS_y$$

(where $dN_i = N(i + 1) - N(i) = 1$) can then be formally replaced by the integral over z

$$v(\boldsymbol{x}) = \frac{I}{4\pi}\int_{-l}^{l} N_o dz \int_{\Sigma(z)} \frac{\partial}{\partial z}\frac{1}{r}\, dS_y \tag{3.76}$$

Case A. Suppose first that the point $\boldsymbol{x}=(X,Y,Z)$ is outside the solenoid. In this case $r(z)>0$ for all $\mathbf{x}$ and z (with $-l\le z\le l$) and, since the geometry of the cross section $\Sigma(z)$ does not depend on z, we have

$$v(\boldsymbol{x})=\frac{I}{4\pi}\int_{-l}^{l}N_o dz\int_{\Sigma(z)}\frac{\partial}{\partial z}\frac{1}{r}\,dS_y=\frac{IN_o}{4\pi}\int_{-l}^{l}dz\frac{\partial}{\partial z}\int_{\Sigma(z)}\frac{1}{r}\,dS$$
$$=\frac{IN_o}{4\pi}\int_{\Sigma(l)}\frac{dS}{r}-\frac{IN_o}{4\pi}\int_{\Sigma(-l)}\frac{dS}{r}$$

Thus outside of the solenoid

$$v(\boldsymbol{x})=\frac{IN_o}{4\pi}\int_{\Sigma(l)}\frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|}-\frac{IN_o}{4\pi}\int_{\Sigma(-l)}\frac{dS_y}{|\boldsymbol{x}-\boldsymbol{y}|}\tag{3.77}$$

is (approximately) the potential of a finite magnetic dipole, that is, of two magnetic single layers distributed on the two bases $\Sigma=\Sigma(-l)$ and $\Sigma'=\Sigma(l)$ at distance $2l$. Comparing with eq. (2.85) shows that the magnetic intensity field $\boldsymbol{H}$ in the exterior of the solenoid coincides with that of a permanent magnet with the same geometry and magnetic moment

$$M_o=IN_o\tag{3.78}$$

in accordance with the Ampère equivalence principle. The outer magnetic induction $\boldsymbol{B}=\mu_o\boldsymbol{H}$ also coincides in the two cases, since we are assuming $\mu=\mu_o$ for both. In particular

$$\boldsymbol{H}\cong\frac{IN_o}{2}\mathbf{c}_3\quad\boldsymbol{B}\cong\mu_o\frac{IN_o}{2}\mathbf{c}_3\tag{3.79}$$

near the solenoid bases.

Case B. Consider now a point $\boldsymbol{x}$ on the solenoid axis, $\boldsymbol{x}=(0,0,Z)$. Eq. (3.76) shows that the magnetic field can be obtained by superposing the contributions (3.70) to the magnetic field [7] due to N_o circular loops per unit length distributed in the interval $(-l,l)$:

$$\boldsymbol{H}(0,0,Z)=\frac{1}{2}IN_oR^2\mathrm{c}_3\int_{-l}^{l}\frac{dz}{(R^2+(Z-z)^2)^{3/2}}$$

[7] superposing first the potentials and then differentiating would require a careful handling of the potential jumps

Since

$$\frac{R^2}{(R^2+(Z-z)^2)^{3/2}} = \frac{d}{dz}\left[\frac{z-Z}{(R^2+(Z-z)^2)^{1/2}}\right]$$

the field on the $Z-$axis, inside and outside of the solenoid, is given by

$$\boldsymbol{H}(0,0,Z) = \frac{1}{2}IN_o\left[\frac{l-Z}{\sqrt{R^2+(l-Z)^2}} + \frac{l+Z}{\sqrt{R^2+(l+Z)^2}}\right]\mathbf{c}_3$$

Under the assumption $l \gg R$ we obtain

$$\boldsymbol{H}(0,0,0) = IN_o\mathbf{c}_3\ ,\ \boldsymbol{H}(0,0,\pm l) = \frac{IN_o}{2}\mathbf{c}_3 \tag{3.80}$$

and from eq. (3.79) we see that $\boldsymbol{H}$ is continuous across the solenoid bases. On the other hand, for $|Z \pm l| \gg R$, i.e. in the central part of the solenoid, $\boldsymbol{H}$ is approximately constant

$$\boldsymbol{H}(0,0,Z) \cong IN_o\,\mathbf{c}_3 \tag{3.81}$$

If the solenoid is thin ($l \gg R$), the values of $\boldsymbol{H}$ and $\boldsymbol{B}$ inside the solenoid can be approximated by those on the axis. We conclude that the magnetic field $\boldsymbol{H}$ and the magnetic induction $\boldsymbol{B}= \mu_o\mathbf{H}$ are approximately constant in the central part of the solenoid [8]

$$\boldsymbol{H} \cong N_oI\mathbf{c}_3\,,\quad \mathbf{B} \cong \mu_oN_oI\mathbf{c}_3 \tag{3.82}$$

and are given by eq. (3.79) near the bases, across which $\boldsymbol{H}\cdot\boldsymbol{n}$ and $\boldsymbol{B}\cdot\boldsymbol{n}$ are continuous.

It is interesting to compare the behavior of $\boldsymbol{B}$ and $\boldsymbol{H}$ in the case of a solenoid and of a permanent magnet with the same geometry and with magnetization $M_o = IN_o$. From eqs (3.79), (3.82) and eqs. (2.88), (2.89) we see that:

(i) the behavior of the induction $\boldsymbol{B}$ is approximately the same, inside and outside the solenoid/magnet

(ii) Outside the solenoid/magnet the behavior of $\boldsymbol{H}$ is approximately the same, and

$$\boldsymbol{H} = \frac{1}{\mu_o}\boldsymbol{B} \tag{3.83}$$

[8] equations (3.82) are exact for $l \to \infty$

(*iii*) For the solenoid **H** is continuous through the bases and satisfies (3.83) everywhere

(*iv*) In the interior of the magnet

$$\boldsymbol{H} \cong \frac{1}{\mu_o}\boldsymbol{B} - IN_o\mathbf{c}_3$$

so that $\boldsymbol{H}$ suffers a discontinuity jump IN_o and a direction reversal at the bases $Z = \pm l$, where $\boldsymbol{B} \cong \frac{1}{2}\mu_o N_o I \mathbf{c}_3$.

The corresponding behavior of the lines of force is depicted in Fig. 3.3 and Fig. 3.4.

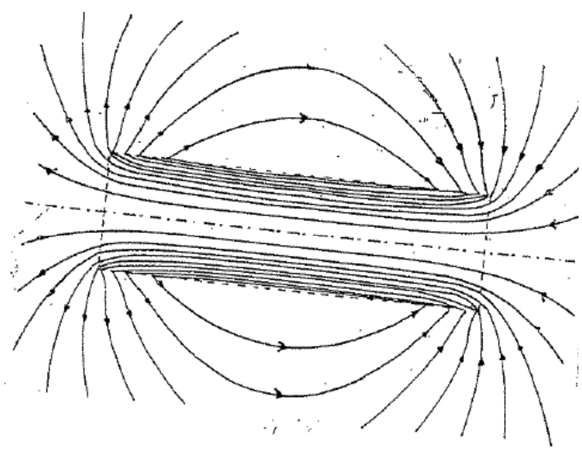

Figure 3.3: Lines of force of $\boldsymbol{H}$ and $\boldsymbol{B}$ for the solenoid and of $\boldsymbol{B}$ for the magnet

The magnetic flux linking the solenoid of length $2l \gg R$ and $N = 2lN_o$ turns is given, under the approximation of eq. (3.82), by

$$\Phi \cong \pi R^2 N \boldsymbol{B} \cdot \mathbf{c}_3 \cong \frac{\pi R^2 I}{2l}\mu_o N^2$$

The solenoid inductance by virtue of eq. (3.33) is then

$$L \cong \frac{\pi R^2 I}{2l}\mu_o N^2 \tag{3.84}$$

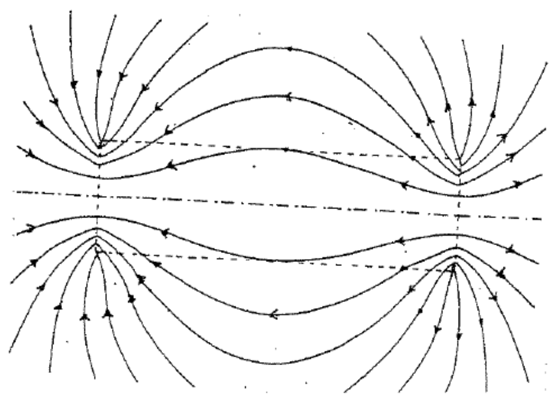

Figure 3.4: Lines of force of $\boldsymbol{H}$ for the magnet

and the total magnetic energy is given by

$$\mathcal{E}_m = \frac{1}{2}LI^2 \cong \frac{1}{2}\mathcal{V}\mu_o N_o^2 I^2 \tag{3.85}$$

where $\mathcal{V} = 2\pi R^2 l$ is the volume of the solenoid and N is the number of turns of the wire. Thus, L and $\mathcal{E}_m$ are proportional to N^2 times the magnetic permeability μ_o of the solenoid core.

Suppose now that, leaving the current unchanged, the solenoid core is replaced by a magnetic material, like soft iron, whose average magnetic permeability is very high, $\mu \cong 10^4\mu_o$. Then inside the soft iron core $\boldsymbol{H}$ does not change, but since $\boldsymbol{B}= \mu\boldsymbol{H}$, eq. (3.82) must be replaced by

$$\boldsymbol{H} \cong N_o I \mathbf{c}_3\,, \qquad \mathbf{B} \cong \mu N_o I \mathbf{c}_3$$

Since $\boldsymbol{B}\cdot\boldsymbol{n}$ is still continuous across the solenoid bases (unlike $\boldsymbol{H}\cdot\boldsymbol{n}$), the magnetic induction is μ/μ_o times higher in modulus outside the solenoid. Similarly,the flux Φ, the self-inductance L and the energy $\mathcal{E}_m$ of the solenoid increase by a factor μ/μ_o due to the insertion of the soft iron core leaving I unchanged. This energy increase takes place at the expense of the external work required initially in order to establish the steady current I in the solenoid.

3.7 Quasi-stationary fields: electric circuits.

In the case of time-dependent fields in conductors the coupling of electric and magnetic fields via the Faraday induction law, the Ampère circuital law and Ohm's law, eqs. (1.23), (1.25) and (C1) of Chapter 1, becomes considerably more complex, involving, as it does, a reciprocal feedback between magnetic fields and self-induced eddy currents. This is true even if the displacement current in conductors is neglected according to the approximation of quasi−stationary fields (§1.7). The situation greatly simplifies in the case of fixed rigid conducting wires of small cross section, since then the currents are almost linear, their direction is prescribed by the geometry of the wires, and the circuital law reduces essentially to the determination of the inductance matrix. For simplicity we restrict our attention here to a single closed $\mathcal{R}L\mathbb{C}$ circuit with concentrated parameters, consisting of a resistor $\mathcal{R}$, a condenser $\mathbb{C}$, an inductor L and a generator which furnishes a given impressed e.m.f. $e_i(t)$. We neglect the resistance, capacity and self-inductance distributed along the wire, and we adopt the hypothesis of quasi−stationary fields. The displacement current is negligible in the conducting wire Γ, but

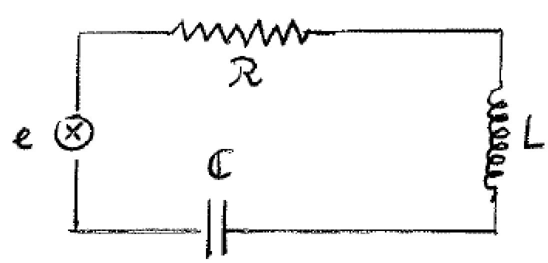

Figure 3.5: $\mathcal{R}L\mathbb{C}$-circuits

not in the space between the condenser plates. In order to see this, consider a closed curve Γ' linking the circuit Γ and take two surfaces S_1, S_2 with common boundary Γ' and such that the first cuts the wire and the second separates the two condenser plates. In this way the displacement current $\frac{\partial \mathbf{D}}{\partial t}$ vanishes on S_1 and the conduction current $\boldsymbol{J}$ vanishes on S_2. Applying the Ampère circuital law yields then

$$\int_{\Gamma'} \boldsymbol{H} \cdot \boldsymbol{t}' \, ds = -\int_{S_2} \frac{\partial \boldsymbol{D}}{\partial t} \cdot \boldsymbol{n} \, dS = \int_{S_1} \boldsymbol{J} \cdot \boldsymbol{n} \, dS := I(t)$$

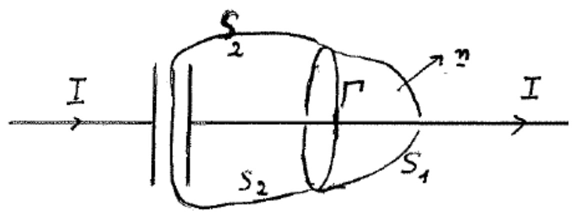

Figure 3.6: Displacement current between the condenser plates

If, at a certain time t, $I(t)$ is different from zero, the integral over S_2 , which represents the displacement current between the two plates, must be different from zero, too. Thus a displacement current must necessarily exist between the two condenser plates, as first recognized by Maxwell, and Γ can be envisaged as a closed circuit, by regarding the dielectric between the condenser plates as part of the circuit.

Note that the assumption of quasi-stationary fields implies that the current $I(t)$ carried by the circuit at time t propagates instantaneously with infinite speed and is the same at all points of Γ (cfr. §1.7). Thus $I(t)$ depends only on time and can take here positive or negative values. In the Faraday-Lenz induction law (3.13) one has to take into account the potential drop due to the condenser. According to eq. (2.76), this potential drop is given by

$$\frac{Q}{\mathbb{C}}$$

where for definiteness $Q = Q(t)$ denotes the charge on that condenser plate towards which the tangent $\boldsymbol{t}$ to Γ is directed. Thus the Faraday induction law takes the form

$$\frac{d\Phi}{dt} = e_{\mathrm{i}} - \mathcal{R}I - \frac{Q}{\mathbb{C}} \tag{3.86}$$

where I is related to the charge Q by

$$I = \frac{dQ}{dt} \tag{3.87}$$

and the Ampere circuital law reduces to eq. (3.72)

$$\Phi(t) = LI(t) \tag{3.88}$$

Putting all these formulae together we arrive at the differential equation

$$L\frac{dI}{dt} + \mathcal{R}I(t) + \frac{1}{\mathbb{C}}Q(t) = e_{\mathrm{i}}(t) \tag{3.89}$$

where the impresses e.m.f. $e_{\mathrm{i}}(t)$ is assigned, and the term LdI/dt can be interpreted as the voltage drop due to the inductor. In a steady state $dI/dt = 0$ and we obtain the relation

$$\mathcal{R}I - e_{\mathrm{i}} + \frac{1}{\mathbb{C}}Q = 0$$

which says that the circulation of $\boldsymbol{E}$ vanishes, in accordance with **H1** of §3.2.2. Note that a steady state implies either $\mathcal{R}I = 0$ or $\mathbb{C} = \infty$, because of eq. (3.87). Substituting (3.87) into eq. (3.89) we obtain the linear differential equation with constant coefficients in the unknown condenser charge $Q(t)$:

$$L\frac{d^2Q}{dt^2} + \mathcal{R}\frac{dQ}{dt} + \frac{1}{\mathbb{C}}Q(t) = e_{\mathrm{i}}(t) \tag{3.90}$$

If $L \neq 0$ this equation is of second order and if we assign the initial charge $Q(0)$ and the initial current $I(0) = dQ(0)/dt$ we obtain the Cauchy problem

$$\begin{aligned} & L\frac{d^2Q}{dt^2} + \mathcal{R}\frac{dQ}{dt} + \frac{1}{\mathbb{C}}Q(t) = e_{\mathrm{i}}(t) \qquad (t > 0) \\ & Q(0) = Q_o \quad , \quad \frac{dQ}{dt}\Big|_{t=0} = I_o \end{aligned}$$

whose unique solution can be written as

$$Q = Q_{tr}(t) + Q_{\infty}(t)$$

where $Q_{tr}(t)$ is the transient, defined as the general integral of the homogeneous equation

$$L\frac{d^2Q}{dt^2} + \mathcal{R}\frac{dQ}{dt} + \frac{1}{\mathbb{C}}Q(t) = 0 \tag{3.91}$$

and the regime term $Q_{\infty}(t)$ is a particular solution of eq. (3.90). Since for any choice of initial data

$$\lim_{t\to+\infty} Q_{tr}(t) = 0 \tag{3.92}$$

(Exercise 7) we have

$$Q(t) \sim Q_\infty(t) \qquad \text{as} \;\; t \to +\infty$$

In practice, $Q(t) \cong Q_\infty(t)$ after a finite relaxation time τ, and the damping of Q_{tr} as $t \to \infty$ is oscillatory if $0 < \mathcal{R} < 2\sqrt{L/\mathbb{C}}$, non-oscillatory if $\mathcal{R} \geq 2\sqrt{L/\mathbb{C}}$

Multiplying eq. (3.90) by $I(t)$ and taking (3.87) into account yields the energy theorem for the circuit

$$\frac{d}{dt}\left(\frac{1}{2}LI^2 + \frac{1}{2}\frac{Q^2}{\mathbb{C}}\right) + \mathcal{R}I^2 = I(t)e_\mathrm{i}(t) \tag{3.93}$$

where, from eq. (3.30) and from eq. (2.52),

$$\frac{1}{2}LI^2 + \frac{1}{2}\frac{Q^2}{\mathbb{C}}$$

is the total energy, the sum of the magnetic energy $\frac{1}{2}LI^2$ and of the electric energy $\frac{1}{2}\frac{Q^2}{\mathbb{C}}$, $\mathcal{R}I^2$ is the power dissipated into heat by the Joule effect (see eq. (3.38)), and $I(t)e_\mathrm{i}(t)$ is the power impressed by the generator. We now consider a few simple examples.

Example 1. **(**$L = 0$**).** The loading of a condenser by means of a constant impressed e.m.f. $e_\mathrm{i}(t) = V_\mathrm{i}$ is described by the Cauchy problem with $L = 0$

$$\mathcal{R}\frac{dQ(t)}{dt} + \frac{1}{\mathbb{C}}Q(t) = V_\mathrm{i} \quad \text{for } t > 0\;; \qquad Q(0) = 0 \tag{3.94}$$

The solution is

$$Q(t) = \mathbb{C}V_\mathrm{i}(1 - e^{-t/\mathcal{R}\mathbb{C}})$$

so that $Q_\infty = \mathbb{C}V_\mathrm{i}$, $Q_{tr} = -\mathbb{C}V_\mathrm{i}e^{-t/\tau}$, and the relaxation time is $\tau = \mathcal{R}\mathbb{C}$. Moreover

$$I(t) = \frac{dQ(t)}{dt} = \frac{V_\mathrm{i}}{\mathcal{R}}\,e^{-t/\mathcal{R}\mathbb{C}} \tag{3.95}$$

Since $Q(t) \to \mathbb{C}V_\mathrm{i} = Q_\infty$ as $t \to +\infty$, the total power P spent in order to build up the final energy $\frac{1}{2}\mathbb{C}V_\mathrm{i}^2$ follows from eq. (3.87):

$$P = \int_0^{+\infty} I(t)e_\mathrm{i}(t)dt = V_\mathrm{i}\int_0^{+\infty}\frac{dQ(t)}{dt}dt = \mathbb{C}V_\mathrm{i}^2$$

and is twice the energy accumulated in the condenser. This is due to the fact that half the power $\mathbb{C}V_{\rm i}^2/2$ is dissipated into heat by the Joule effect:

$$\int_0^\infty \mathcal{R}I^2 dt = \frac{1}{2}\mathbb{C}V_{\rm i}^2$$

Example 2. ($\mathbb{C} = \infty$)**.** The loading of an inductor, i.e. the building up of a magnetic field $H = N_o I$ in a solenoid (see eq. (3.82)) in the absence of the condenser, is described by the Cauchy problem

$$L\frac{dI}{dt} + \mathcal{R}I(t) = V_{\rm i}, \quad I(0) = 0$$

The solution is

$$I(t) = \frac{V_{\rm i}}{\mathcal{R}}(1 - e^{-\mathcal{R}t/L}) \tag{3.96}$$

Here the relaxation time is $\tau = L/\mathcal{R}$ and $I(t) \to I_\infty := V_{\rm i}/\mathcal{R}$ as $t \to +\infty$, but this limit involves an infinite power expenditure due to an infinite Joule dissipation, as can easily be seen from eq. (3.96):almost all the impressed power is dissipated into heat, and the process must in practice be interrupted after a long but finite time $t \gg \tau$. On the other hand, the theoretical power expenditure due solely to the complete loading of the inductor, subtracting the Joule dissipation, is given by

$$\int_0^\infty \left[I\,V_{\rm i} - \mathcal{R}I^2\right]dt = \frac{V_{\rm i}^2}{\mathcal{R}}\int_0^\infty (e^{-t/\tau} - e^{-2t/\tau})\,dt = \frac{1}{2}LI_\infty^2$$

In any case, for given $V_{\rm i}$ and $\mathcal{R}$ the power expenditure is proportional to L, and hence to the magnetic permeability μ of the solenoid core.

Example 3. (simple a-c series circuit). An alternating current circuit has a periodic forcing term $e_{\rm i}(t)$ which is typically of the form

$$e_{\rm i}(t) = e_o sin\,\omega t$$

Then $I(t) = I_\infty(t)$+terms vanishing as $t \to +\infty$, where the regime solution $I_\infty(t)$ is also periodic of period $2\pi/\omega$, with amplitude I_o and phase-shift δ depending on the (circular) frequency ω :

$$I_\infty(t) = I_o\,sin\,(\omega t - \delta) \quad , \qquad I_o = \frac{e_o}{Z(\omega)} \quad , \quad \delta = arcos\,\frac{\mathcal{R}}{Z(\omega)}$$

Here $cos\,\delta = \mathcal{R}/Z(\omega)$ is the power factor and $Z = Z(\omega)$ the impedance

$$Z = \sqrt{\mathcal{R}^2 + (\omega L - \frac{1}{\omega\,\mathbb{C}})^2}$$

Thus the amplitude $I_o(\omega)$ is maximum (in absolute value) at resonance [36], that is for

$$\omega = \omega_{\text{res}} := \frac{1}{\sqrt{L\mathbb{C}}}$$

and correspondingly $Z(\omega_{\text{res}}) = \mathcal{R}$ is minimum and $\delta(\omega_{\text{res}}) = 0$. The mean square power dissipated in the a-c circuit by the Joule effect is

$$\frac{1}{2} e_o I_o cos\,\delta \;=\; \frac{1}{2}\, I_o^2 \mathcal{R} \;=\; \overline{I^2}\,\mathcal{R}$$

where the bar denotes time average over one period $2\pi/\omega$ (see Exercise 8).

Example 4 (**ideal transformer**). An ideal transformer consists of a primary solenoid carrying the current I_1 and consisting of n_1 turns of wire of cross-section area A_1 and length l, inductively coupled with a secondary solenoid carrying the current $I_2(t)$ and consisting of n_2 turns of wire of cross-section area A_2 and length l. The inductive coupling is ideal, in the sense that all the magnetic field is confined in the transformer core, the lines of force Γ are closed, and the total m.m.f. is zero

$$\oint_\Gamma \boldsymbol{H} \cdot \boldsymbol{t}\, ds = (n_1 I_1 - n_2 I_2) = 0$$

(see Fig. 3.7) Thus if $\kappa := n_1/n_2$ denotes the turns ratio we have

$$I_2 = \kappa I_1 \tag{3.97}$$

The voltage drops due to the inductors (neglecting the mutual inductances) are

$$e_1(t) = L_{11}\frac{dI_1}{dt} \quad , \quad e_2(t) = L_{22}\frac{dI_2}{dt} \tag{3.98}$$

and from eq. (3.84) with the same L and $\mathcal{R}$ we have $L_{11} \propto n_1^2$, $L_{22} \propto n_2^2$ and $L_{11} = \kappa^2 L_{22}$. Combining eqs. (3.97) and (3.98) yields

$$e_1(t) = \kappa^2 L_{22}\frac{dI_1}{dt} \quad , \quad e_2(t) = \kappa L_{22}\frac{dI_1}{dt}$$

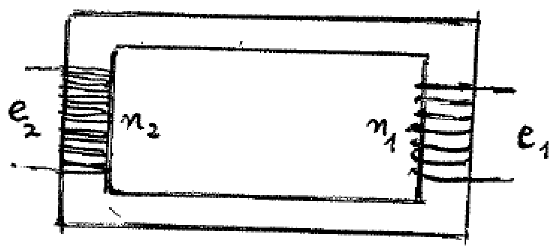

Figure 3.7: Ideal transformer

so that

$$e_1(t) = \kappa e_2(t)$$

In other words, the ratio of the primary and secondary voltages is equal to the ratio κ of the number of turns in their windings, while the ratio of the currents is the inverse ratio κ^{-1} and the power is invariant, $e_1 I_1 = e_2 I_2$.

Ohm's law $e_j = I_j \mathcal{R}_j = I_j \frac{n_j l}{\gamma A_j} \propto I_j \frac{n_j}{A_j}$, $j = 1, 2$ (see eq. (1.36)) implies that the wire cross-sections areas must be in the ratio

$$\frac{A_2}{A_1} = \frac{n_2 I_2}{e_2} \frac{e_1}{n_1 I_1} = \kappa$$

Thus the higher-voltage winding will have more turns of wire of smaller cross-section.

Exercises

Exercise 1. Choosing another closed irreducible curve $\tilde{\Gamma}$ and applying the Stokes' theorem yields

$$\oint_{\Gamma} \boldsymbol{J} \cdot \boldsymbol{t}\, ds - \oint_{\tilde{\Gamma}} \boldsymbol{J} \cdot t\, ds = \int_{\Sigma} curl\, \boldsymbol{J} \cdot \boldsymbol{n}\ dS = 0$$

as $curl\, \boldsymbol{J} = \boldsymbol{0}$ in $\mathcal{T}$. Here Σ denotes a surface strip contained in $\mathcal{T}$ and bounded by $\Gamma_{\rm i}$ and $\tilde{\Gamma}$. Thus the circulation depends only on the homology class of Γ [11].

Exercise 2. Eq. (3.23) yields

$$|\boldsymbol{H}(\boldsymbol{x})| \le \frac{1}{4\pi}\int_{\mathcal{T}} \frac{|\boldsymbol{J}(\boldsymbol{y})|}{|\boldsymbol{x}-\boldsymbol{y}|^{2}}\, \boldsymbol{dy} \ , \quad \mathbf{x}\in\mathbb{R}^3$$

where $\mathcal{T}$ is bounded, so that

$$|\boldsymbol{x}-\boldsymbol{y}|^{-2} = (|\boldsymbol{x}|^2+|\boldsymbol{y}|^2-2\boldsymbol{x}\cdot\boldsymbol{y})^{-1} = |\boldsymbol{x}|^{-2}\,(1+\frac{|\boldsymbol{y}|^2-2\boldsymbol{x}\cdot\boldsymbol{y}}{|\boldsymbol{x}|^2})^{-1} \sim |\boldsymbol{x}|^{-2}$$

as $|\boldsymbol{x}| \to +\infty$. Thus

$$|\boldsymbol{H}(\boldsymbol{x})| \le \frac{1}{4\pi|\boldsymbol{x}|^{2}}\int_{\mathcal{T}} |\boldsymbol{J}(\boldsymbol{y})|\, \boldsymbol{dy} \ , \quad \boldsymbol{x}\in\mathbb{R}^3$$

Excersize 3. For a wire we have $J \cong \frac{I}{A}$,hence

$$\frac{1}{\gamma I^2}\int_{\text{wire}} |\boldsymbol{J}(\boldsymbol{y})|^2 dV \cong \frac{l\,I^2}{\gamma\,A\,I^2} \cong \frac{l}{\gamma A}$$

Exercise 4. By choosing polar coordinates $(r\,,\varphi)$ we have

$$v(0,0,z) = \frac{I}{4\pi}\int_{\Sigma} \frac{\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{y})}{|\boldsymbol{x}-\boldsymbol{y}|^3} dS_y = \frac{2\pi I z}{4\pi}\int_0^R \frac{r\,dr}{(z^2+r^2)^{3/2}}$$

since $\boldsymbol{n}=\mathbf{c}_3$ and $\mathbf{c}_3\cdot(\boldsymbol{x}-\boldsymbol{y}) = z$. Performing the integration yields

$$v(0,0,z) = \frac{I}{2}(\frac{z}{|z|} - \frac{z}{\sqrt{R^2+z^2}}) \quad \to \begin{cases} I/2 & z\to 0+ \\ -I/2 & z\to 0- \end{cases}$$

and differentiating we obtain the magnetic field

$$\boldsymbol{H}(0,0,z) = \frac{I\ R^2\ \mathbf{c}_3}{2(R^2+z^2)^{3/2}}$$

which is smooth at all points of the loop axis.

Exercise 5. Show that the curvilinear double integral

$$\mathcal{I} = \int_{\Gamma}\int_{\Gamma} \frac{ds_x ds_y}{|\boldsymbol{x}-\boldsymbol{y}|}$$

is divergent.

Exercise 6. The self-induced action (resultant force) in the case of a toroidal conductor is zero. Hint:

$$\begin{aligned}\boldsymbol{F} &= \mu_o \int_{\mathcal{T}} \boldsymbol{J} \wedge \boldsymbol{H} dV = \frac{\mu_o}{4\pi} \int_{\mathcal{T}} \int_{\mathcal{T}} \boldsymbol{J}(\boldsymbol{x}) \wedge (grad_x \frac{1}{r} \wedge \boldsymbol{J}(\boldsymbol{y}))\mathrm{dxdy} \\ &= \frac{\mu_o}{4\pi} \int_{\mathcal{T}} \int_{\mathcal{T}} \boldsymbol{J}(\boldsymbol{x}) \cdot \boldsymbol{J}(\boldsymbol{y}) grad_x \frac{1}{r}\, \mathrm{dxdy} - \frac{\mu_o}{4\pi} \int_{\mathcal{T}} \int_{\mathcal{T}} (\boldsymbol{J}(\boldsymbol{x}) \cdot grad_x \frac{1}{r}) \boldsymbol{J}(\boldsymbol{y}) \mathrm{dxdy} \\ &= -\frac{\mu_o}{4\pi} \int_{\mathcal{T}} \int_{\mathcal{T}} div_x(\frac{1}{r} \boldsymbol{J}(\boldsymbol{x}) \otimes \boldsymbol{J}(\boldsymbol{y}))\mathrm{dxdy} \;= \boldsymbol{0} \qquad (r = |\boldsymbol{x} - \boldsymbol{y}|\,)\end{aligned}$$

since $div\, \boldsymbol{J} = 0$ in $\mathcal{T}$, $\boldsymbol{J} \cdot \boldsymbol{n} = 0$ on $\partial \mathcal{T}$ and the first integral on the second line is zero by symmetry.

Exercise 7. The difference of two regime solutions Q_∞ satisfies the homogeneous equation (3.91) because of linearity.

(i) If $\mathcal{R}/2L > 1/\sqrt{L\mathbb{C}}$ the general integral of (3.91)

$$Q_{tr}(t) = e^{-\mathcal{R}t/2L} \Big[A\, exp\,(\sqrt{(\frac{\mathcal{R}}{2L})^2 - \frac{1}{L\mathbb{C}}}\, t) + Bexp(-\sqrt{(\frac{\mathcal{R}}{2L})^2 - \frac{1}{L\mathbb{C}}} t)\Big]$$

is a monotone damped function of t.

(ii) If $0 < \mathcal{R}/2L < 1/\sqrt{L\mathbb{C}}$ we have the damped oscillations

$$Q_{tr}(t) = e^{-\mathcal{R}t/2L} \Big[Acos\,(\sqrt{\frac{1}{L\mathbb{C}} - (\frac{\mathcal{R}}{2L})^2}\, t) + Bsin(\sqrt{\frac{1}{L\mathbb{C}} - (\frac{\mathcal{R}}{2L})^2} t)\Big]$$

(iii) Finally if $\mathcal{R}/2L = 1/\sqrt{L\mathbb{C}}$ the general integral

$$Q_{tr}(t) = e^{-t/\sqrt{L\mathbb{C}}}(A + Bt)$$

is monotone for t large enough and the damping rate is larger than in case (i). The constants A and B are fixed by the initial data.

Exercise 8. If $\mathcal{R} > 0$ and $e_{\mathrm{i}}(t) = e_o sin\omega t$ the regime solution for $Q(t)$ is given by

$$Q_\infty = A\, cos\,(\omega t - \delta) + \text{ terms vanishing as } t \to +\infty$$

where the amplitude $A = A(\omega)$ is given by

$$A(\omega) = -\frac{e_o}{\omega Z}$$

and the phase-shift $\delta = \delta(\omega)$ is

$$\delta(\omega) = arctan\frac{X_L - X_{\mathbb{C}}}{\mathcal{R}} \equiv arctan\frac{\omega L - (\omega\mathbb{C})^{-1}}{\mathcal{R}} \equiv arcos\frac{\mathcal{R}}{Z(\omega)}$$

$X_L := \omega L, X_{\mathbb{C}} := \frac{1}{\omega\,\mathbb{C}}$ are called the reactances and

$$Z(\omega) := \sqrt{\mathcal{R}^2 + (X_L - X_{\mathbb{C}})^2} = \sqrt{\mathcal{R}^2 + (\omega L - \frac{1}{\omega\,\mathbb{C}})^2} \ = \frac{\mathcal{R}}{cos\,\delta}$$

is called the impedance of the a-c circuit. It follows that the regime current dQ_∞/dt is

$$I_\infty(t) = I_o\,sin\,(\omega t - \delta(\omega)) \ + \text{ terms vanishing as } t \to +\infty$$

with amplitude

$$I_o(\omega) = \frac{e_o}{Z(\omega)} = \frac{e_o cos\,\delta(\omega)}{\mathcal{R}}$$

Since

$$I^2_\infty(t) = I^2_o\,sin^2(\omega t - \delta)\ , \quad I^2_o = \frac{e_o I_o}{Z} \ = \frac{e_o I_o}{\mathcal{R}}\,cos\,\delta(\omega)$$

averaging $I(t) = I_\infty(t)$ over a period yields

$$\mathcal{R}\overline{I^2} = I^2_o(\omega)\,\frac{\mathcal{R}\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} sin^2(\omega t - \delta)\,dt = \frac{1}{2}\mathcal{R}I^2_o(\omega) \ = \frac{1}{2}I_o e_o\frac{\mathcal{R}}{Z} = \frac{1}{2}\,e_o I_o cos\,\delta(\omega)$$

At resonance $\omega_{res} = 1/\sqrt{L\mathbb{C}}$, $X_L = X_{\mathbb{C}}$, $Z(\omega)$ attains its minimum value $Z(\frac{1}{\sqrt{L\mathbb{C}}}) = \mathcal{R}$, $I_o(\omega)$ attains its maximum absolute value

$$I_o(\omega_{\rm res}) \ = \frac{e_o}{Z(\omega)} \ = \frac{e_o}{\mathcal{R}}$$

and the phase shift $\delta(\omega_{\rm res}) = arcos(1) = 0$, so that the current is in phase with the e.m.f. $e_{\rm i}(t)$.

If $\mathcal{R} = 0$ and $\omega \neq \omega_{\rm res}$ the previous solution still holds, with $\delta(\omega) = \pi/2$ and $Z(\omega) = |X_L - X_{\mathbb{C}}| > 0$.

If $\mathcal{R} = 0$ and $\omega = 1/\sqrt{L\mathbb{C}}$ the impedance vanishes and the regime solution Q_∞ at resonance is no longer periodic but contains the secular term

$$\frac{1}{2}\sqrt{\frac{\mathbb{C}}{L}}\,t\,cos\,\omega t$$

Exercise 9. Prove that for a system of two parallel straigth wire conductors with length l and distance d, the self-inductance and the mutual inductance are given by

$$L_{11} = L_{22} \cong \frac{\mu_o l}{2\pi}\, log\, \frac{2l}{r_o} \quad , \quad L_{12} \cong \frac{\mu_o l}{2\pi}\, log\, \frac{2l}{d}$$

where r_o is the radius of the circular cross-section of the wires, and $l \gg d \gg r_o$.

Chapter 4

Electromagnetic Waves

This chapter is devoted to a study of the full time-dependent Maxwell equations for homogeneous non-magnetic media when the displacement current cannot be neglected and the quasi-stationary approximation for low frequency phenomena cannot be applied. The electromagnetic field can then be represented in terms of a scalar potential u and a vector potential $\boldsymbol{V}$, determined up to a "gauge transformation" (§4.4).

The full Maxwell system is intimately related to the wave equation and hence to wave propagation phenomena with a finite speed (§4.1). Of particular interest are the plane monochromatic waves, discussed in §4.2. Like the wave equation, the Maxwell system admits characteristic surfaces, that can be interpreted as wavefronts in ordinary space, and bicharacteristic rays that play an important role in the high frequency approximation known as geometrical optics, briefly discussed in §4.7.

The Cauchy problem for the Maxwell equations in a homogeneous medium is dealt with in §4.3 and §4.4 as regards classical solutions, and in §4.6 as regards weak solutions. Explicit representation formulae for classical solutions are given via the method of spherical means and Kirchhoff's retarded potentials.

After deriving the laws of reflection and refraction at the boundary between two different media (Snell's law), evanescent waves and the phenomenon of total reflection are discussed in §4.8. In §4.9 and §4.10 we deal with the problem of wave propagation through a one-dimensional layered medium, at normal or oblique incidence, formulated in precise mathematical terms as a transmission problem for the Maxwell equations in one space variable. Explicit exact and approximate solutions are given, and the interesting phenomenon of reflection reduction in the periodic case is discussed.

From a mathematical point of view, the unifying thread of this chapter is hyperbolicity. The full Maxwell system is hyperbolic and even the telegraph equation, discussed in §4.5, though obtained by neglecting the displacement current, turns out to be a variant of the wave equation.

4.1 Maxwell's equations and wave propagation

We recall that the Maxwell equations for homogeneous non-magnetic media are

$$\mu \frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E}\,, \quad div\, \boldsymbol{H} = 0 \tag{4.1}$$

$$\epsilon \frac{\partial \boldsymbol{E}}{\partial t} = curl\, \boldsymbol{H} - \boldsymbol{J}\,, \quad div \boldsymbol{E} = \rho/\epsilon \tag{4.2}$$

where $\boldsymbol{J} = \gamma \boldsymbol{E}$ satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + div \boldsymbol{J} = 0 \tag{4.3}$$

which follows from eq. (4.2), and ϵ, μ, γ are constants such that

$$\epsilon \geq \epsilon_o > 0 \quad , \quad \mu > 0 \quad , \quad \gamma \geq 0$$

Therefore the first two equations (4.1), (4.2) can be reduced to the normal form

$$\frac{\partial \boldsymbol{H}}{\partial t} = -\frac{1}{\mu}\, curl\, \boldsymbol{E}\;, \qquad \frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{\epsilon} curl\, \boldsymbol{H} - \frac{\gamma}{\epsilon}\, \boldsymbol{E} \tag{4.4}$$

If $(\boldsymbol{E}, \boldsymbol{H})$ is a solution of class C^2, these equations have an important consequence, first discovered by Maxwell. Differentiating the second equation with respect to t, taking the *curl* of the first, applying Schwartz's theorem and eliminating $\boldsymbol{H}$ yields the damped vector wave equation for $\boldsymbol{E}$

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} + \frac{\gamma}{\epsilon}\frac{\partial \boldsymbol{E}}{\partial t} = -c^2 curl\, curl\, \boldsymbol{E}$$

where the constant c is defined by

$$c := (\epsilon\mu)^{-1/2} \tag{4.5}$$

The vector identity (1.76), valid for cartesian components, yields

$$curl curl\, \boldsymbol{E} \;\equiv\; grad\, div\, \boldsymbol{E} - \Delta_3 \boldsymbol{E} \tag{4.6}$$

If $\rho \equiv 0$ we have $div\, \boldsymbol{E} = 0$ and this identity implies

$$curl curl\, \boldsymbol{E} \;\equiv -\Delta_3 \boldsymbol{E}$$

Thus if $\rho \equiv 0$ each cartesian component E_k of the electric field $\boldsymbol{E}$ in a conductor with conductivity γ satisfies the damped wave equation

$$\frac{\partial^2 E_k}{\partial t^2} + \frac{\gamma}{\epsilon}\frac{\partial E_k}{\partial t} = c^2\, \Delta_3 E_k \qquad (k = 1, 2, 3) \tag{4.7}$$

which in a dielectric ($\gamma = 0$) reduces to the ordinary homogeneous wave equation

$$\frac{\partial^2 E_k}{\partial t^2} = c^2 \Delta_3 E_k \qquad (k = 1, 2, 3)$$

The same homogeneous wave equation holds for the cartesian components of $\boldsymbol{H}$ and hence of $\boldsymbol{D} = \epsilon \boldsymbol{E}$ and $\boldsymbol{B} = \mu \boldsymbol{H}$.

It is important to remark that this crucial fact is due to the presence of the displacement current $\partial \boldsymbol{D}/\partial t$, one of the fundamental ideas of Maxwell. We have seen in §1.7 that neglecting the displacement current would yield a different partial differential equation, namely the heat equation, which governs phenomena evolving in time with nominally infinite speed [2]. In contrast, the wave equation obtained here describes propagation phenomena with finite speed c defined by eq. (4.5). But, as first realized by Maxwell, this numerical value for c coincides with the speed of light:for example, in empty space $c = c_o$, where

$$c_o := (\epsilon_o \mu_o)^{-1/2} \cong 301,000 \text{ km/sec.} \tag{4.8}$$

is the same as the speed of propagation of light in vacuo, already known from experiments in Maxwell's time. In a dielectric medium with permittivity ϵ the speed of light (4.5) can be written as

$$c = c_o/n_r$$

where n_r is the refractive index

$$n_r := \sqrt{\epsilon/\epsilon_o} \tag{4.9}$$

In conclusion, the electromagnetic waves propagate in an isotropic medium with the speed of light c given by (4.5), light is an electromagnetic wave $(\boldsymbol{E},\boldsymbol{H})$, which is able to propagate even in empty space with speed c_o, and Optics becomes a chapter of Electromagnetism. [1]

4.2 Plane waves

An important class of waves is represented by linearly polarized plane monochromatic waves, which can be written in the complex form [2]

$$\boldsymbol{E} = \boldsymbol{E}_o e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})} \quad , \quad \boldsymbol{H} = \boldsymbol{H}_o e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})} \tag{4.10}$$

where $\omega > 0$ is the (circular) frequency, $\boldsymbol{p}$ is the wavenumber vector, and $\boldsymbol{E}_o$, $\boldsymbol{H}_o$ are the amplitudes. As we will see, $\boldsymbol{p}$, $\boldsymbol{E}_o$ and $\boldsymbol{H}_o$ are in general complex vectors. In general, a wave is called plane if the surfaces of constant phase are planes; monocromatic if it is characterized by a single frequency ω; and linearly polarized if the directions of the electric and magnetic vectors are fixed in space. A monochromatic wave is obviously time-periodic, with period $2\pi/\omega$, and is a particular case of a traveling wave.

The importance of plane monochromatic waves lies also in the fact that an arbitrary linearly polarized wave can always be represented as their superposition by means of a Fourier series or integral.

[1] The electromagnetic waves were first revealed by Hertz (1888).

[2] the use of complex numbers greatly simplifies the calculations, at least if linear expressions in $\boldsymbol{E}$ and $\boldsymbol{H}$ are considered, as we do in this chapter. It is understood that, at the end, one must take the real part of the results

4.2.1 Homogeneous and heterogeneous waves.

Let us look for solutions of the Maxwell equations of the form (4.10). Since the time dependence is expressed by the complex exponential $e^{i\omega t}$, eqs. (4.4) take the form

$$curl\,\boldsymbol{E} = -i\omega\mu\boldsymbol{H}\ , \quad curl\,\boldsymbol{H} = \gamma\,\boldsymbol{E} + i\omega\epsilon\boldsymbol{E}\ \equiv\ i\omega\epsilon'\boldsymbol{E} \tag{4.11}$$

where ϵ' denotes the complex permittivity

$$\epsilon' := \epsilon - i\frac{\gamma}{\omega} \tag{4.12}$$

Equations (4.11), called harmonic Maxwell equations, imply that $\boldsymbol{E}$ and $\boldsymbol{H}$ are solenoidal

$$div\,\boldsymbol{E}\ =\ div\,\boldsymbol{H}\ = 0$$

and therefore we must have

$$\rho \equiv 0$$

(Time-harmonic solutions of the Maxwell equations do not exist unless ρ vanishes identically, as can also be seen by adapting the proof of Proposition 1.3.1.) Since the plane wave (4.10) depends upon the space variable $\boldsymbol{x}$ via the complex exponential $e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}$, it is easily verified (Exercise 1) that

$$div\,\boldsymbol{E} = -i\boldsymbol{p}\cdot\boldsymbol{E}\quad ,\quad curl\,\boldsymbol{E} = -i\boldsymbol{p}\wedge\boldsymbol{E}$$

and similarly for $\boldsymbol{H}$. Eqs. (4.11) and (4.13) thus reduce to the vector relations

$$\boldsymbol{p}\cdot\boldsymbol{E}_o = 0\ ,\ \ \boldsymbol{p}\cdot\boldsymbol{H}_o = 0 \tag{4.13}$$

and

$$\boldsymbol{H} = \frac{1}{\omega\mu}\boldsymbol{p}\wedge\boldsymbol{E}\ ,\quad \boldsymbol{E} = \frac{1}{\omega\epsilon'}\boldsymbol{H}\wedge\boldsymbol{p} \tag{4.14}$$

These equations show that $\boldsymbol{E}_o$, $\boldsymbol{H}_o$ and $\boldsymbol{p}$ are three mutually orthogonal vectors. Being complex vectors, however, this orthogonality property has no immediate geometrical meaning, unless all the three vectors are real, which happens only if $\epsilon' = \epsilon$ is real, that is for $\gamma = 0$. This corresponds to propagation in a dielectric medium.

I. Plane monochromatic waves in dielectrics: $\gamma = 0$, $\epsilon' = \epsilon$.

Since ϵ' is real we assume here that $\boldsymbol{E}_o$, $\boldsymbol{H}_o$ and $\boldsymbol{p}$ are real vectors. The wavenumber vector

$$\boldsymbol{p} = p\,\mathbf{k}\ ;\quad p := |\boldsymbol{p}|\ ,\quad |\mathbf{k}| = 1$$

which yields the direction of propagation of the wave, is orthogonal to $\boldsymbol{E}$ and $\boldsymbol{H}$ by force of (4.13). In other words, the vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ lie in a plane transversal to the direction of propagation, so that the wave is transversal. The quantity $\omega t - \boldsymbol{p}\cdot\boldsymbol{x}$ is known as the phase of the wave, and every level surface, or surface of constant phase

$$\omega t - \boldsymbol{p}\cdot\boldsymbol{x} = \text{constant}$$

is a plane orthogonal to $\mathbf{k}$ which moves in space with normal velocity $\mathbf{v}_f = v_f\mathbf{k}$, called phase velocity of the wave. Since $\omega t - \boldsymbol{p}\cdot\boldsymbol{x}$=constant implies

$$\omega dt - p\mathbf{k}\cdot d\boldsymbol{x} = 0 \quad \Rightarrow \quad v_f = \mathbf{k}\cdot d\boldsymbol{x}/dt = \omega/p$$

the phase velocity is given here by

$$v_f := \frac{\omega}{p}$$

Eliminating $\boldsymbol{E}$ from eqs. (4.14) yields, taking (4.13) into account,

$$\boldsymbol{H} = \frac{1}{\omega^2\epsilon'\mu}\ \boldsymbol{p}\wedge(\boldsymbol{H}\wedge\boldsymbol{p}) \equiv \frac{1}{\omega^2\epsilon'\mu}\ (\boldsymbol{p}\cdot\boldsymbol{p}\,\boldsymbol{H} - \boldsymbol{p}\cdot\boldsymbol{H}\,\boldsymbol{p}) \equiv \frac{\boldsymbol{H}\ p^2}{\omega^2\epsilon'\mu}$$

and for $\boldsymbol{H}\neq\mathbf{0}$ this implies

$$\boldsymbol{p}\cdot\boldsymbol{p} = \omega^2\epsilon'\mu \tag{4.15}$$

(a similar results is obtained by eliminating $\boldsymbol{H}$). Since $\gamma = 0$, $\epsilon' = \epsilon$ it follows that the wavenumber is related to the frequency by the dispersion relation

$$p = p(\omega) \equiv \frac{\omega}{c}$$

where c is given by (4.5). If we assign $\omega, \mathbf{k}$ and $\boldsymbol{E}_o$ arbitrarily, with $\omega > 0$, $\mathbf{k}\in\mathbb{R}^3$, $\boldsymbol{E}_o\in\mathbb{R}^3$, $|\mathbf{k}| = 1$, $\mathbf{k}\cdot\boldsymbol{E}_o = 0$, we have

$$\boldsymbol{p} = p(\omega)\mathbf{k} = \frac{\omega}{c}\mathbf{k}\,,\ \boldsymbol{H}_o = \frac{1}{\omega\mu}\ \boldsymbol{p}\wedge\boldsymbol{E}_o \equiv \frac{1}{\mu c}\mathbf{k}\wedge\boldsymbol{E}_o \tag{4.16}$$

and the plane monochromatic wave (obtained by taking the real parts)

$$(4.17)\qquad \boldsymbol{E}=\boldsymbol{E}_o cos(\omega(t\,-\,\mathbf{k}\cdot\boldsymbol{x}/c)\ ,\ \boldsymbol{H}=\frac{1}{\mu c}\mathbf{k}\wedge\boldsymbol{E}_o\,cos(\omega(t\,-\,\mathbf{k}\cdot\boldsymbol{x}/c)$$

is periodic in t with period $2\pi/\omega$ and wavelength

$$\lambda := 2\pi/p = 2\pi c/\omega$$

Since the dispersion relation $p=p(\omega)$ is linear and homogeneous, the phase velocity coincides with the group velocity, defined by

$$\mathrm{v}_g = v_g\mathbf{k},\ \ v_g := \frac{d\omega}{dp}\equiv\frac{1}{dp(\omega)/d\omega}$$

and both are given by

$$\mathbf{v}_f=\mathrm{v}_g=c\mathbf{k}\equiv(\epsilon\mu)^{-1/2}\mathbf{k}$$

The refractive index, defined by

$$(4.18)\qquad n_r := \frac{c_o}{v_f}\equiv\frac{c_o}{c}$$

(cfr. (4.9)) is thus independent of ω :a plane wave with this property is called non-dispersive.

From the preceding relations we find that the electric and magnetic fields of the wave satisfy

$$(4.19)\qquad \boldsymbol{H}=\sqrt{\frac{\epsilon}{\mu}}\ \mathbf{k}\wedge\boldsymbol{E},\quad \boldsymbol{E}=\sqrt{\frac{\mu}{\epsilon}}\boldsymbol{H}\wedge\mathbf{k}$$

where the quantity $\sqrt{\frac{\mu}{\epsilon}}$ is the wave impedance, and the Poynting vector

$$\boldsymbol{S}:=\boldsymbol{E}\wedge\boldsymbol{H}=\frac{1}{\mu c}\mathbf{k}|\boldsymbol{E}_o|^2cos^2(\omega(t\,-\,\mathbf{k}\cdot\boldsymbol{x}/c))$$

is parallel to $\boldsymbol{p}= p\mathbf{k}$ and to $\boldsymbol{c}:= c\mathbf{k}$. Moreover, the definition (4.5) of c shows that

$$(4.20)\qquad \frac{1}{2}\epsilon|\boldsymbol{E}|^2=\frac{1}{2}\mu|\boldsymbol{H}|^2$$

and

$$\boldsymbol{S} = \epsilon|\boldsymbol{E}|^2 c\mathbf{k} \;\equiv\; \mu|\boldsymbol{H}|^2 c\mathbf{k} \;\equiv\; W\mathbf{c}$$

where $W = \frac{1}{2}\epsilon|\boldsymbol{E}|^2 + \frac{1}{2}\mu|\boldsymbol{H}|^2 \equiv W_e + W_m$, see §1.5. Thus $W_e = W_m$ by virtue of eq. (4.20) (equipartition of energy), and the specific power flux radiated by the plane wave is given by

$$\boldsymbol{S} = W\mathbf{c} \equiv 2W_e\mathbf{c} \equiv 2W_m\mathbf{c}$$

To summarize: For any choice of frequency ω, direction $\mathbf{k}$ and magnetic amplitude $\boldsymbol{H}_o$ (with $\mathbf{k}\cdot\boldsymbol{H}_o = 0$, $|\mathbf{k}| = 1$), a dielectric supports linearly polarized, transverse, non-dispersive plane waves propagating with velocity $c := (\epsilon\mu)^{-1/2}$, wavenumber $\boldsymbol{p} = \frac{\omega}{c}\mathbf{k}$, and electric amplitude $\boldsymbol{E}_o = \sqrt{\frac{\mu}{\epsilon}}\, \boldsymbol{H}_o \wedge \mathbf{k}$.

II. Plane monochromatic waves in conductors: $\gamma > 0$.

We consider only the case of low or moderate frequencies ω, satisfying the restriction

$$\omega \ll \frac{\gamma}{\epsilon} \tag{4.21}$$

The complex permittivity (4.12) is then pure imaginary

$$\epsilon' \cong -i\frac{\gamma}{\omega}$$

and eqs. (4.14), (4.15) become

$$\boldsymbol{H} = \frac{1}{\omega\mu}\boldsymbol{p} \wedge \boldsymbol{E} \quad , \quad \boldsymbol{E} = \frac{i}{\gamma}\,\boldsymbol{H} \wedge \boldsymbol{p} \tag{4.22}$$

whence

$$\boldsymbol{p} \cdot \boldsymbol{p} = -\, i\mu\gamma\,\omega \tag{4.23}$$

Therefore the wavenumber $\boldsymbol{p}$ cannot be real, but must have the form

$$\boldsymbol{p} = \boldsymbol{P}' - i\boldsymbol{p}'$$

with $\boldsymbol{p}', \boldsymbol{P}' \in \mathbb{R}^3$. Eq. (4.23) becomes then

$$|\boldsymbol{P}'|^2 - |\boldsymbol{p}'|^2 - 2i\boldsymbol{P}' \cdot \boldsymbol{p}' = -\, i\mu\gamma\omega \tag{4.24}$$

and by equating real and imaginary parts we obtain

$$|\boldsymbol{P}'| = |\boldsymbol{p}'| \ , \quad \boldsymbol{P}' \cdot \boldsymbol{p}' = \frac{1}{2}\gamma\mu\omega > 0 \tag{4.25}$$

The second equation shows that the vectors $\boldsymbol{P}'$ and $\boldsymbol{p}'$ cannot be mutually orthogonal. The most interesting case is when they coincide:

$$\boldsymbol{P}' = \boldsymbol{p}' = p\mathbf{k} \qquad \Rightarrow \qquad \boldsymbol{p} = (1-i)p\,\mathbf{k}$$

The dispersion relation is then

$$p = \, p(\omega) \doteq \sqrt{\frac{\gamma\mu\omega}{2}}$$

and since this relation is nonlinear, the phase velocity and the group velocity are different:

$$v_f \ := \frac{\omega}{p} = \frac{\sqrt{2\omega}}{\sqrt{\gamma\mu}} \ , \quad v_g := \frac{d\omega}{dp} = 2v_f$$

As $\omega \ll \gamma/\epsilon$ we have $v_g = 2v_f \ll c$ and the index of refraction

$$n_r = \frac{c_o}{v_f} = c_o\sqrt{\frac{\gamma\,\mu}{2\omega}}$$

is a function of frequency. Thus the wave is dispersive: conductors are dispersive media, as anticipated §1.3.1. The wavenumber is

$$\boldsymbol{p} = \sqrt{\gamma\mu\omega}\,\frac{1-i}{\sqrt{2}}\mathbf{k} \ \equiv \ e^{-i\pi/4}\sqrt{\gamma\mu\omega}\,\mathbf{k}$$

and eqs. (4.22) take the form

$$\boldsymbol{H} = e^{-i\pi/4}\sqrt{\frac{\gamma}{\mu\omega}}\ \mathbf{k}\wedge\boldsymbol{E} \quad , \quad \boldsymbol{E} = e^{i\pi/4}\sqrt{\frac{\mu\omega}{\gamma}}\ \boldsymbol{H}\wedge\mathbf{k} \tag{4.26}$$

If the amplitude vector $\boldsymbol{H}_o$ (say) is real it follows that $\boldsymbol{E}_o$ is complex, but the vector $e^{-i\pi/4}\boldsymbol{E}= \sqrt{\frac{\mu\omega}{\gamma}}\,\boldsymbol{H}\wedge\mathbf{k}$ is also real and is a given function of ω, $\boldsymbol{H}_o$ and $\mathbf{k}$. By choosing a coordinate system such that $\mathbf{k}=\mathbf{c}_3$, the (real part of) the plane wave (4.26) reads

$$\boldsymbol{E} = \boldsymbol{E}_o\, e^{-z/\delta} cos(\omega t - \frac{z}{\delta(\omega)} + \frac{\pi}{4})\,, \boldsymbol{H} = \boldsymbol{H}_o\, e^{-z/\delta}\, cos(\omega t - \frac{z}{\delta(\omega)}) \tag{4.27}$$

where $z = \mathbf{c}_3 \cdot \boldsymbol{x}$, $\boldsymbol{H}_o$ is an arbitrary real vector orthogonal to $\mathbf{c}_3$, $\boldsymbol{E}_o = \sqrt{\frac{\mu\omega}{\gamma}}\ \boldsymbol{H}_o \wedge \mathbf{c}_3$, and $\delta(\omega)$ is the inverse of the scalar wavenumber $p(\omega)$:

$$\delta := \sqrt{\frac{2}{\gamma\mu\omega}}$$

already introduced in eq. (1.79). Note that the waves (4.27) are damped in the direction of propagation $\mathbf{k}$, due to the presence of the damping term in eq. (4.7):conductors are absorbing media, and every dispersive medium is also absorbing (cfr. §1.7).

To summarize: For frequencies ω satisfying (4.21) and for any choice of $\mathbf{k}$ and $\boldsymbol{H}_o$ (with $\mathbf{k} \cdot \boldsymbol{H}_o = 0$, $|\mathbf{k}| = 1$), a conductor supports linearly polarized, transverse, dispersive, damped plane waves (4.27) propagating with speed $v_f = \omega\delta \ll c := (\epsilon\mu)^{-1/2}$, wavenumber

$$\boldsymbol{p} = \sqrt{\gamma\mu\,\omega}\,\frac{1-i}{\sqrt{2}}\,\mathbf{k}$$

and electric field with amplitude $\boldsymbol{E}_o = \sqrt{\frac{\mu\omega}{\gamma}} e^{i\pi/4} \boldsymbol{H}_o \wedge \mathbf{k}$. The surfaces of constant phase $z =$constant coincide with the surfaces of constant amplitude for these waves, called homogeneous plane waves. An analysis of eq. (4.27) shows that they satisfy the following three Fourier laws [3]:

(1) phase shift: the field vectors $\boldsymbol{E}$, $\boldsymbol{H}$, and $\boldsymbol{J} = \gamma\,\boldsymbol{E}$ for $z > 0$ have a phase delay with respect to $z = 0$ equal to z/δ for $\boldsymbol{H}$ and $z/\delta - \pi/4$ for $\boldsymbol{J}$ and $\boldsymbol{E}$;

(2) damping: $\boldsymbol{E}$**,** $\boldsymbol{H}$**,** $\boldsymbol{J}$ are damped exponentially for increasing $z = \mathbf{k} \cdot \boldsymbol{x}$ in the conductor for fixed t with damping rate $\delta^{-1}(\omega)$, due to the presence of the imaginary wavenumber $\boldsymbol{p}'$;

(3) skin effect:the damping rate $\delta^{-1}(\omega)$ increases with increasing frequency (precisely, with increasing $\sqrt{\gamma\omega}$), and $\delta^{-1}(\omega) \to \infty$ as $\omega \to \infty$.

These laws, verified experimentally, are also valid in the quasi-stationary approximation, which consists in neglecting the displacement current with respect to the conduction current, and in fact condition (4.21) coincides with (1.80) of §1.7.

[3] first found by Fourier for solutions of the Earth temperature problem for the heat equation [2]

Remark 1. The case, which will not be considered here, in which $\boldsymbol{p}' \neq \boldsymbol{P}'$ and the two vectors are not parallel characterizes the so-called heterogeneous plane waves in conductors [35]. Such waves are damped, dispersive and their constant phase surfaces, which are planes orthogonal to ***P'***, are distinct from the constant amplitude surfaces, which are planes orthogonal to $\boldsymbol{p}'$.

4.2.2 Evanescent waves in dielectrics.

Heterogeneous plane damped waves with complex wavenumber $\boldsymbol{p}=\boldsymbol{P}' - i\boldsymbol{p}'$

$$\boldsymbol{E} = \boldsymbol{E}_o e^{-\,\boldsymbol{p}'\cdot \boldsymbol{x}}\, e^{i(\omega t - \boldsymbol{P}'\cdot \boldsymbol{x})}, \qquad \boldsymbol{H} = \boldsymbol{H}_o e^{-\,\boldsymbol{p}'\cdot \boldsymbol{x}}\, e^{i(\omega t - \boldsymbol{P}'\cdot \boldsymbol{x})}$$

are possible also in dielectrics, with $\gamma = 0$ and $\epsilon' = \epsilon$, provided $\boldsymbol{p}'$ and $\boldsymbol{P}'$ are orthogonal vectors, so that the damping is orthogonal to the propagation. Indeed, if

$$\gamma = 0 \ \ , \ \epsilon' = \epsilon \ \ , \quad \mathbf{p} = \mathbf{P}' - i\mathbf{p}'$$

eq. (4.15) becomes

$$|\boldsymbol{P}'|^2 - |\boldsymbol{p}'|^2 - 2i\boldsymbol{P}'\cdot\boldsymbol{p}' \ = \ \omega^2 \epsilon\, \mu$$

and if

$$\boldsymbol{P}'\cdot\boldsymbol{p}' \ = \ 0 \tag{4.28}$$

this equation can be solved in the form $\boldsymbol{P}' \ = \ P'(\omega)\mathbf{k}$, where $\mathbf{k}$ is an arbitrary real unit vector and

$$P'(\,\omega) = \ \sqrt{|\boldsymbol{p}'|^2 \ + \frac{\omega^2}{c^2}} \tag{4.29}$$

which plays the role of a dispersion relation, provided $\boldsymbol{p}' = \boldsymbol{p}'(\omega)$ is known. The phase velocity and the refractive index are given by eq. (4.29) in the form (Exercise 2)

$$v_f := \frac{\omega}{P'} \ = \frac{c}{\sqrt{1 + |\boldsymbol{p}'|^2 \frac{c^2}{\omega^2}}} \ \ , \ \ n_r := \frac{c_o}{v_f} = \frac{c_o}{c}\sqrt{1 + |\boldsymbol{p}'|^2 \frac{c^2}{\omega^2}}$$

and if $\boldsymbol{p}'(\omega) \propto \omega$ the wave is non-dispersive. These waves are called evanescent, since they decay exponentially (and often very fast) as $\boldsymbol{p}'\cdot\boldsymbol{x} > 0$ increases [4]. The evanescent waves are in general not transversal, and the surfaces of constant phase $\boldsymbol{P}'\cdot\boldsymbol{x} - \omega t$ =constant are orthogonal to those of constant amplitude, $\boldsymbol{p}'\cdot\boldsymbol{x}$=constant.

A typical example arises from the phenomenon of reflection and refraction at a surface in optics. At grazing incidence the wavenumber becomes complex and the transmitted wave becomes an evanescent wave which disappears completely in the high frequency limit of geometrical optics. In this example $\boldsymbol{p}'$ is determined by Snell's law and is proportional to ω, so that n_r does not depend on ω, and the evanescent transmitted wave is non-dispersive.

Remark 2. Evanescent waves are impossible in a conductor, since eq. (4.25) implies that $\boldsymbol{P}'$ and $\boldsymbol{p}'$ cannot be orthogonal for $\gamma > 0$ (and $\omega > 0$).

4.2.3 Wavegroups.

An arbitrary linearly polarized electric field $\boldsymbol{E}(\boldsymbol{x}, t)$ or magnetic field $\boldsymbol{H}(\boldsymbol{x}, t)$ can be represented, under assumptions well known from Calculus, as a superposition of plane monochromatic waves by means of a Fourier series expansion

$$\boldsymbol{E}(\boldsymbol{x}, t) = \sum_{n=-\infty}^{+\infty} e^{i(n\omega t - \boldsymbol{p}(n\omega)\cdot\boldsymbol{x})} \boldsymbol{E}_o(n\omega)$$

or a Fourier integral

$$\boldsymbol{E}(\boldsymbol{x}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i(\omega t - \boldsymbol{p}(\omega)\cdot\boldsymbol{x})} \boldsymbol{E}_o(\omega)\, d\omega \tag{4.30}$$

where $\boldsymbol{E}_o(\omega)$ is the amplitude vector and, if $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy the Maxwell equations, $\boldsymbol{p}(\omega)$ is given by the dispersion relation. In other words, an arbitrary plane non-monochromatic wave can be represented as a group of monochromatic plane waves traveling in space. If $\boldsymbol{p}(\omega) = p(\omega)\mathbf{k}$ the velocity of the wavegroup (4.30) is by definition the group velocity $\boldsymbol{v}_g$.

Definition 4.2.1 *A wavegroup (4.30) is called a wavepacket if* $\boldsymbol{E}_o(\omega)$ *has compact support concentrated around a frequency* ω_o *. The group velocity* $\mathbf{v}_g$

[4] for this reasons, evanescent waves are also called surface waves

of a wavepacket with $\boldsymbol{p}(\omega) = p(\omega)\boldsymbol{k}$ is the velocity with which the variation in the shape of the wave envelope (also called amplitude modulation) propagates in physical space.

In some cases (when $v_g < c$) we can also envisage the group velocity as the velocity of propagation of the signal. To clarify this definition, we give here two examples.

Example 1. In the absence of dispersion, $\boldsymbol{p}(\omega) = \frac{\omega}{c}\mathbf{k}, \boldsymbol{E}_o(\omega) = E_o(\omega)\boldsymbol{\kappa}\,,\ \boldsymbol{\kappa}\cdot\mathbf{k} = 0$ and (4.30) yields the traveling wave

$$\boldsymbol{E}(\boldsymbol{x},t) = \frac{\boldsymbol{\kappa}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega(t-\mathbf{k}\cdot\boldsymbol{x}/c)}E_o(\omega)\,d\omega \;=\; e_o(t\;-\;\mathbf{k}\cdot\frac{\boldsymbol{x}}{c}\,)\,\boldsymbol{\kappa}$$

where $e_o(t\,) := \mathcal{F}^{-1}[E_o]$ is the inverse Fourier transform of $E_o(\omega)$ [2]. This wave travels without distortion with the group velocity $\mathbf{v}_g = c\mathbf{k}$,which coincides with the phase velocity $\mathbf{v}_f$ and is independent of the frequency ω.

In the absence of dispersion, all the Fourier components travel with the same phase velocity $\mathbf{v}_f$ which coincides with the velocity $\mathbf{v}_g$ of the wavegroup in physical space.

Example 2. Suppose that $\boldsymbol{p}$=$\boldsymbol{p}(\omega)\mathbf{k}$, $\boldsymbol{E}_o(\omega) = E_o(\omega)\boldsymbol{\kappa}\,,\ \boldsymbol{\kappa}\cdot\mathbf{k} = 0$ and that the dispersion relation $p = p(\omega)$ is non linear (dispersive case). Let $E_o(\omega) = G(\omega-\omega_o)$ be a positive impulse with support $[\omega_o-\varepsilon, \omega_o-\varepsilon]$ sharply concentrated at a frequency ω_o corresponding to the maximum for $|E_o(\omega)|$. We can then write [5]

$$\boldsymbol{p}(\omega) \cong p_o\mathbf{k} + (\omega-\omega_o)p_o'\,\mathbf{k}\ , \quad p_o := p(\omega_o), \quad p_o' = p'(\omega_o)$$

Substituting into (4.30), and setting $\omega = \omega_o + y$, we obtain

$$\begin{aligned} \text{(4.31)}\qquad \boldsymbol{E}(\boldsymbol{x},t) &= \frac{\boldsymbol{\kappa}}{\sqrt{2\pi}}\int_{\omega_o-\varepsilon}^{\omega_o+\varepsilon} e^{i(\omega t-\boldsymbol{p}(\omega)\cdot\boldsymbol{x})}E_o(\omega)d\omega \\ &\cong \frac{\boldsymbol{\kappa}}{\sqrt{2\pi}}\int_{-\varepsilon}^{+\varepsilon} e^{\,i(\omega_o t+yt-p_o\mathbf{k}\cdot\boldsymbol{x}-p_o'y\,\mathbf{k}\cdot\boldsymbol{x})}G(y)dy \\ &= \frac{\boldsymbol{\kappa}}{\sqrt{2\pi}}e^{\,i(\omega_o t-p_o\mathbf{k}\cdot\boldsymbol{x})}\int_{-\infty}^{+\infty} e^{\,iy(t-p_o'\mathbf{k}\cdot\boldsymbol{x})}G(y)\,dy \\ &= e^{\,i(\omega_o t-p_o\mathbf{k}\cdot\boldsymbol{x})}g(t-p_o'\mathbf{k}\cdot\boldsymbol{x})\,\boldsymbol{\kappa} \end{aligned}$$

[5] see V. Cantoni, Introduzione all' equazione di Schrödinger e alle equazioni d'onda relativistiche, unpublished Lecture Notes , Univ. di Milano, Italy, 1993

where $g := \mathcal{F}^{-1}[G]$ is the inverse Fourier transform of G. The superposition (4.31) of individual wavelets of different wavelengths traveling at different (phase) speeds can therefore be envisaged as a monochromatic wave $e^{\,i(\omega_o t - p_o \mathbf{k}\cdot\boldsymbol{x})}$ which travels with the phase velocity

$$\mathbf{v}_f = \frac{\omega_o}{p_o}\mathbf{k}$$

and is amplitude modulated by the function $g(t - p_0'\mathbf{k}\cdot\boldsymbol{x})$. By the Riemann-Lebesgue Lemma [39] $g(t - p_0'\mathbf{k}\cdot\boldsymbol{x})$ and hence $\boldsymbol{E}(\boldsymbol{x}, t)$ tend to zero as $|\boldsymbol{x}|$ and t tend to infinity. The result is therefore a wavepacket localized in time and space and traveling with the group velocity

$$\mathbf{v}_g = \frac{1}{p_0'}\mathbf{k} \;\equiv \frac{d\omega}{dp_o}\,\mathbf{k} \tag{4.32}$$

This velocity is different from the phase velocity $v_f\mathbf{k}$ which characterizes the Fourier component having the maximum amplitude $|E_o(\omega_o)|$ in the Fourier space: in the presence of dispersion the speeds of propagation of the maximum amplitude in the Fourier and physical space are in general different. In several cases the relation

$$v_f\, v_g = c^2$$

is satisfied, so that if $v_f < c$ then $v_g > c$ (cfr. Exercise 2).

4.3 Cauchy problem for the Maxwell equations in vacuo.

We consider the inhomogeneous Maxwell equations in empty space

$$\mu_o \frac{\partial \boldsymbol{H}}{\partial t} \;=\; -curl\,\boldsymbol{E}\,,\;\; \epsilon_o \frac{\partial \boldsymbol{E}}{\partial t} \;= curl\,\boldsymbol{H} - \boldsymbol{J}_r\,, \tag{4.33}$$

$$div\,\boldsymbol{E} \;= \rho/\epsilon_o \;\;,\;\; div\,\boldsymbol{H} = 0$$

for $t > 0$, $\boldsymbol{x} \in \mathbb{R}^3$, with a given current $\boldsymbol{J} = \boldsymbol{J}_r(\boldsymbol{x}, t)$ and initial conditions for $t = 0$

$$\boldsymbol{E}(\boldsymbol{x}, 0) = \boldsymbol{E}_o(\boldsymbol{x})\;,\;\; \boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{H}_o(\boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{R}^3) \tag{4.34}$$

Here the vector functions $\boldsymbol{E}_o(\boldsymbol{x}), \boldsymbol{H}_o(\boldsymbol{x}),\ \boldsymbol{J}_r(\boldsymbol{x},t)$ are assigned , with

$$div\,\boldsymbol{E}_o(\boldsymbol{x}) = 0,\ div\,\boldsymbol{H}_o(\boldsymbol{x}) = 0 \tag{4.35}$$

and for all $\boldsymbol{x}\in \mathbb{R}^3, t > 0$ the charge density $\rho(\boldsymbol{x},t)$ is initially zero and satisfies the continuity equation:

$$\rho(\boldsymbol{x},0) = 0, \qquad \frac{\partial\rho}{\partial t} + div\,\boldsymbol{J}_r(\boldsymbol{x},t) \ = \ 0 \ \ (t > 0) \tag{4.36}$$

Thus ρ is identically zero if $\boldsymbol{J}_r$ is.

This initial value problem is an idealized mathematical model of a radiating antenna K driven by a given current $\boldsymbol{J}_r(\boldsymbol{x},t)$. We have seen in Chapter 3 that for a conductor in the presence of an impressed current Ohm's law becomes $\boldsymbol{J}= \gamma(\boldsymbol{E}+\boldsymbol{E}_{\rm imp})$, and that $\boldsymbol{E}$ and $\boldsymbol{J}$ are known in the conductor in steady conditions if the impressed field $\boldsymbol{E}_{\rm imp}$ is known. We are assuming here that the same is true in the unsteady case, namely that both the impressed current $\boldsymbol{J}_{\rm imp} = \gamma\boldsymbol{E}_{\rm imp}$ and the electric field $\boldsymbol{E}$ are known inside the conductor, so that $\boldsymbol{J}=\boldsymbol{J}_r(\boldsymbol{x},t)$ becomes a known vector function of $\boldsymbol{x}$ and t satisfying (4.36).

Theorem 4.3.1 *(uniqueness). For any given $\boldsymbol{J}_r(\boldsymbol{x},t), \boldsymbol{E}_o(\boldsymbol{x})$ and $\boldsymbol{H}_o(\boldsymbol{x})$ the Cauchy problem (4.33)–(4.35) has at most one solution $(\boldsymbol{E},\boldsymbol{H}) \in C^o(\mathbb{R}^3 \times [0,+\infty)) \cap C^1(\mathbb{R}^3 \times (0,+\infty))$.*

Proof. We need to prove that the homogeneous problem, with zero data $\boldsymbol{J}_r \equiv \boldsymbol{E}_o \equiv \boldsymbol{H}_o \equiv \rho \equiv 0$, has only the trivial solution $\boldsymbol{E}(\boldsymbol{x},t) \equiv \boldsymbol{H}(\boldsymbol{x},t) \equiv 0$ for $\boldsymbol{x}\in \mathbb{R}^3$, $t \geq 0$.

Let $(\boldsymbol{E},\boldsymbol{H})$ be a solution of the homogeneous problem, and let

$$\mathcal{E}[\mathcal{B}(t)] = \frac{1}{2}\int_{\mathcal{B}(t)} (\epsilon_o|\boldsymbol{E}|^2+\mu_o|\boldsymbol{H}|^2)\,d\boldsymbol{x} \ \equiv \frac{1}{2}\int_0^{R-c_ot} d\rho \int_{|\boldsymbol{x}-\boldsymbol{x}_o|=\rho} (\epsilon_o|\boldsymbol{E}|^2+\mu_o|\boldsymbol{H}|^2)\,dS$$

be the corresponding energy in a ball $\mathcal{B}(t) = \big\{(\boldsymbol{x},t) : |\boldsymbol{x}-\boldsymbol{x}_o| \leq R - c_o t\big\}$

contracting with speed $c_o = (\epsilon_o\mu_o)^{-\frac{1}{2}}$. Eqs. (4.33) yield for $t > 0$

$$\frac{d\mathcal{E}[\mathcal{B}(t)]}{dt} = \int_{\mathcal{B}(t)} (\epsilon_o \boldsymbol{E}\cdot\boldsymbol{E}_t + \mu_o \boldsymbol{H}\cdot\boldsymbol{H}_t)\, d\boldsymbol{x} - \frac{1}{2}c_o \int_{\partial\mathcal{B}(t)} (\epsilon_o|\boldsymbol{E}|^2 + \mu_o|\boldsymbol{H}|^2)\, dS$$

$$= -\frac{1}{2}c_o \int_{\partial\mathcal{B}(t)} (\epsilon_o|\boldsymbol{E}|^2 + \mu_o|\boldsymbol{H}|^2)\, dS + \int_{\mathcal{B}(t)} (\boldsymbol{E}\cdot curl\, \boldsymbol{H} - \boldsymbol{H}\cdot curl\, \boldsymbol{E})\, d\boldsymbol{x}$$

$$= -\frac{1}{2} \int_{\partial\mathcal{B}(t)} (\sqrt{\frac{\epsilon_o}{\mu_o}}|\boldsymbol{E}|^2 + \sqrt{\frac{\mu_o}{\epsilon_o}}|\boldsymbol{H}|^2)\, dS + \int_{\mathcal{B}(t)} (\boldsymbol{E}\cdot curl\, \boldsymbol{H} - \boldsymbol{H}\cdot curl\, \boldsymbol{E})\, d\boldsymbol{x}$$

where $\partial\mathcal{B}(t) = \{\boldsymbol{x}\in\mathbb{R}^3 : |\boldsymbol{x}-\boldsymbol{x}_o| = R - c_o t\}$, and from identity (1.48)

$$\int_{\mathcal{B}(t)} (\boldsymbol{E}\cdot curl\, \boldsymbol{H} - \boldsymbol{H}\cdot curl\, \boldsymbol{E})\, d\boldsymbol{x} = \int_{\partial\mathcal{B}(t)} \boldsymbol{H}\wedge\boldsymbol{E}\cdot\boldsymbol{n}\, dS \le \int_{\partial\mathcal{B}(t)} |\boldsymbol{E}||\boldsymbol{H}|\, dS$$

We have then

$$\frac{d\mathcal{E}[\mathcal{B}(t)]}{dt} \le -\frac{1}{2}\int_{\partial\mathcal{B}(t)} \Big[\sqrt[4]{\frac{\epsilon_o}{\mu_o}}|\boldsymbol{E}| - \sqrt[4]{\frac{\mu_o}{\epsilon_o}}|\boldsymbol{H}|\Big]^2 dS \le 0$$

and, proceeding as in the proof of Theorem of §1.5.5, we obtain

$$\mathcal{E}[\mathcal{B}(t)] \le \mathcal{E}[\mathcal{B}(0)] \quad \text{for all } t \ge 0$$

Since $\mathcal{E}[\mathcal{B}(0)] = 0$, this inequality implies $\mathcal{E}[\mathcal{B}(t)] \equiv 0$ for all $t \ge 0$. Thus $\boldsymbol{E}\equiv\boldsymbol{H}\equiv 0$ in $\mathcal{B}(t)$, and the assertion follows letting $R\to\infty$.

By linearity, the solution, if it exists, can be written as the sum

$$\boldsymbol{E}(\boldsymbol{x},t) = \boldsymbol{E}'(\boldsymbol{x},t) + \boldsymbol{E}''(\boldsymbol{x},t), \quad \boldsymbol{H}(\boldsymbol{x},t) = \boldsymbol{H}'(\boldsymbol{x},t) + \boldsymbol{H}''(\boldsymbol{x},t)$$

where $(\boldsymbol{E}',\boldsymbol{H}')$ satisfy the Cauchy problem with $\boldsymbol{J}_r \equiv \boldsymbol{0}$ and $(\boldsymbol{E}'',\boldsymbol{H}'')$ satisfy the problem with $\boldsymbol{E}_o(\boldsymbol{x}) \equiv \boldsymbol{H}_o(\boldsymbol{x}) \equiv \boldsymbol{0}$. In other words, we have

$$\text{(4.37)} \qquad \epsilon_o\frac{\partial\boldsymbol{E}'}{\partial t} = curl\, \boldsymbol{H}',\ \mu_o\frac{\partial\boldsymbol{H}'}{\partial t} = -curl\, \boldsymbol{E}'$$

$$div\, \boldsymbol{E}' = div\, \boldsymbol{H}' = 0 \ (\boldsymbol{x}\in\mathbb{R}^3, t > 0)$$

$$\text{(4.38)} \qquad \boldsymbol{E}'(\boldsymbol{x},0) = \boldsymbol{E}_o(\boldsymbol{x})\ ,\ \boldsymbol{H}'(\boldsymbol{x},0) = \boldsymbol{H}_o(\boldsymbol{x}) \ \ (\boldsymbol{x}\in\mathbb{R}^3)$$

and

$$\mu_o\frac{\partial\boldsymbol{H}''}{\partial t} = -curl\, \boldsymbol{E}'',\ \epsilon_o\frac{\partial\boldsymbol{E}''}{\partial t} = curl\, \boldsymbol{H}'' - \boldsymbol{J}_r,\ div\, \boldsymbol{E}'' = div\, \boldsymbol{H}'' = 0$$

$$\boldsymbol{E}''(\boldsymbol{x},0) = \boldsymbol{0}\ ,\ \boldsymbol{H}''(\boldsymbol{x},0) = \boldsymbol{0} \ \ (\boldsymbol{x}\in\mathbb{R}^3)$$

4.3.1 Spherical means.

A representation formula for the solution ($\boldsymbol{E}'$,$\boldsymbol{H}'$) of the Cauchy problem (4.37) and (4.38), with solenoidal initial data $\boldsymbol{E}_o, \boldsymbol{H}_o$, can be written in terms of spherical means.

Definition 4.3.2 *(spherical means). Let $f(\boldsymbol{x}) \in C^2(\mathbb{R}^3)$. The spherical mean of f over the sphere Σ_r: $|\boldsymbol{y}-\boldsymbol{x}| = r$ with center $\boldsymbol{x}$ and radius $r > 0$ is the function $M = M(\boldsymbol{x}, r) = Mf\{\boldsymbol{x}, r\}$ defined by*

$$Mf\{\boldsymbol{x}, r\} = \frac{1}{4\pi\, r^2} \int_{\Sigma_r} f(\boldsymbol{y})\, dS_y \equiv \frac{1}{4\pi} \int_{\Omega} f(\boldsymbol{x} + r\boldsymbol{\nu})\, d\Omega \tag{4.39}$$

where $\boldsymbol{\nu} = (\boldsymbol{y}-\boldsymbol{x})/|\boldsymbol{y}-\boldsymbol{x}|$ is the normal to the unit sphere Ω and $d\Omega$ the element of solid angle in $\mathbb{R}^3$.

Since the integral (4.39) does not change by changing the sign of $\boldsymbol{\nu}$, the spherical mean can be extended to $r < 0$ as an even function

$$M(\boldsymbol{x}, -r) = M(\boldsymbol{x}, r)$$

and for $r = 0$ we will set, by definition,

$$M(\boldsymbol{x}, 0\} := f(\boldsymbol{x}) \quad , \; M_r(\boldsymbol{x}, 0) := 0 \; , \;\; M_{rr}(\boldsymbol{x}, 0) := \frac{1}{3}\, \Delta_3 f(\boldsymbol{x}) \tag{4.40}$$

(M_r is the partial derivative of M with respect to r, etc.). With this position it is easy to see that $M(\boldsymbol{x}, r)$ has the following properties (Exercise 4):

(i) $\lim_{r\to 0} Mf\{\boldsymbol{x}, r\} = f(\boldsymbol{x})$

(ii) $\lim_{r\to 0} \frac{\partial}{\partial r}(Mf\{\boldsymbol{x}, r\}) = 0$

(iii) $\lim_{r\to 0} \frac{\partial^2}{\partial r^2}(Mf\{\boldsymbol{x}, r\}) = \frac{1}{3}\, \Delta_3 f(\boldsymbol{x})$

(iv) $M(\boldsymbol{x}, r) \in C^2(\mathbb{R}^4)$

(v) $M\Delta_3 f \equiv \Delta_3 Mf$.

(vi) $M(\boldsymbol{x}, r)$ satisfies the Darboux equation in all $\mathbb{R}^4$ (Δ_3 is the Laplacian with respect to $\boldsymbol{x}$):

$$M_{rr} + \frac{2}{r} M_r = \Delta_3\, M$$

The left-hand side of the Darboux equation is the radial Laplacian in 3D, which satisfies the identity

$$M_{rr} + \frac{2}{r} M_{rr} \equiv \frac{1}{r}(rM)_{rr}$$

It follows that the Darboux equation, multiplied by r, can also be written in the form

$$(rM)_{rr} = r\Delta_3 M \equiv \Delta_3(rM) \tag{4.41}$$

so that by the change of variable $r = c_o t$ we obtain the wave equation for the function $c_o t M\{\boldsymbol{x}, c_o t\}$.

Theorem 4.3.3 *(a) If $f(\boldsymbol{x}) \in C^2(\mathbb{R}^3)$, the scalar spherical mean $tM = tMf\{\boldsymbol{x}, c_o t\}$ is a solution of class $C^2(\mathbb{R}^4)$ of the wave equation*

$$\frac{\partial^2 (tM)}{\partial t^2} = c_o^2 \, \Delta_3 \, (tM) \tag{4.42}$$

and satisfies the initial conditions

$$tM\Big|_{t=0} = 0 \, , \quad \frac{\partial \, (tM)}{\partial t}\Big|_{t=0} = f(\boldsymbol{x})$$

(b) If $\boldsymbol{v}(\boldsymbol{x}) \in C^2(\mathbb{R}^3)$ is a solenoidal vector field, the vector spherical mean $t\boldsymbol{M} = tM\boldsymbol{v}\{\boldsymbol{x}, c_o t\}$ in cartesian coordinates is a $C^2(\mathbb{R}^4)$ solution of the vector wave equation

$$\frac{\partial^2 (t\boldsymbol{M})}{\partial t^2} = c_o^2 \, \Delta_3 (t\boldsymbol{M}) \equiv -c_o^2 \, curl \, curl (t\boldsymbol{M}) \tag{4.43}$$

Proof. (a) Eq. (4.42) coincides with eq. (4.41) written for $r = c_o t$, apart from c_o factors, which can be eliminated. Property (i) implies that tM vanishes for $t = 0$, and,since

$$(tM)_t \equiv M + tM_t$$

from properties (i), (ii) we find that $(tM)_t = f(\boldsymbol{x})$ for $t = 0$.

(b) The following chain of identities holds for $t\boldsymbol{M} = tM\boldsymbol{v}\{\boldsymbol{x}, c_o t\}$:

$$curl \, curl (t\boldsymbol{M}) = t \, curl \, curl \boldsymbol{M} = tM curl \, curl \boldsymbol{v} = -tM\Delta_3 \boldsymbol{v} = -\Delta_3 (tM\boldsymbol{v})$$

by force of the identity (4.6) applied to $\boldsymbol{v}$, with $div \, \boldsymbol{v} = 0$.

4.3.2 Loss of derivatives. Huygens' principle.

Theorem 4.3.3 enables us to solve the Cauchy problem for the homogeneous Maxwell equations (4.37), (4.38).

Theorem 4.3.4 *Let $\boldsymbol{E}_o, \boldsymbol{H}_o \in C^2(\mathbb{R}^3)$ satisfy (4.35). The Cauchy problem (4.37), (4.38) has a unique solution $(\boldsymbol{E}', \boldsymbol{H}')$ in $C^1(\overline{D})$, $D = \mathbb{R}^3 \times (0, +\infty)$ which can be represented in terms of the vector spherical means of $\boldsymbol{E}_o, \boldsymbol{H}_o$:*

$$\boldsymbol{E}'(\boldsymbol{x},t) = \frac{t}{\epsilon_o} M\, curl\, \boldsymbol{H}_o + \frac{\partial}{\partial t}(tM\boldsymbol{E}_o) \tag{4.44}$$

$$\boldsymbol{H}'(\boldsymbol{x},t) = -\frac{t}{\mu_o} M\, curl\, \boldsymbol{E}_o + \frac{\partial}{\partial t}(tM\boldsymbol{H}_o)$$

Proof. By property (iv) $\boldsymbol{E}'$ and $\boldsymbol{H}'$ are in $C^1(\overline{D})$. Properties $(i), (ii)$ and eq. (4.40) imply that

$$\lim_{t\to 0} \frac{\partial}{\partial t}(t\, M\boldsymbol{E}_o) \equiv \lim_{t\to 0}(M\boldsymbol{E}_o + t\frac{\partial}{\partial t}M\boldsymbol{E}_o) = \boldsymbol{E}_o(\boldsymbol{x})$$

and similarly for $\frac{\partial}{\partial t}(t\, M\boldsymbol{H}_o)$. Therefore the initial conditions (4.38) are satisfied. Taking the *curl* of $\boldsymbol{E}'$ in the first equation (4.44) yields

$$\begin{aligned} curl\, \boldsymbol{E}' &= \frac{1}{\epsilon_o} tM curl\, curl \boldsymbol{H}_o + \frac{\partial}{\partial t}(tM curl \boldsymbol{E}_o) \\ &= -\frac{1}{\epsilon_o} tM\Delta_3 \boldsymbol{H}_o + \mu_o \frac{\partial^2}{\partial t^2}(tM\boldsymbol{H}_o) - \mu_o \frac{\partial \boldsymbol{H}'}{\partial t} \end{aligned}$$

Since $c_o = (\epsilon_o \mu_o)^{-\frac{1}{2}}$, eq. (4.43) implies that

$$\mu_o \frac{\partial^2}{\partial t^2}(tM\boldsymbol{H}_o) = \mu_o c_o^2 \Delta_3 (tM\boldsymbol{H}_o) \equiv \frac{1}{\epsilon_o} tM\Delta_3 \boldsymbol{H}_o$$

It follows that

$$curl\, \boldsymbol{E}' = -\mu_o \frac{\partial \boldsymbol{H}'}{\partial t}$$

Similarly, taking the *curl* of $\boldsymbol{H}'$ in the second equation (4.44) shows that

$$curl\, \boldsymbol{H}' = \epsilon_o \frac{\partial \boldsymbol{E}'}{\partial t}$$

Moreover, from eq.(4.44) we find

$$div\, \boldsymbol{E}' = \frac{1}{\epsilon_o} tdiv\, M curl \boldsymbol{H}_o + \frac{\partial}{\partial t}(tdiv\, M \boldsymbol{E}_o) \equiv \frac{1}{\epsilon_o} tM div\, curl \boldsymbol{H}_o + \frac{\partial}{\partial t}(tM div\, \boldsymbol{E}_o) = 0$$

and

$$\begin{aligned} div\, \boldsymbol{H}' &= -\frac{1}{\mu_o} tdiv\, M curl\, \boldsymbol{E}_o + \frac{\partial}{\partial t}(tdiv\, M \boldsymbol{H}_o) \\ &\equiv -\frac{1}{\mu_o} tM div\, curl \boldsymbol{E}_o + \frac{\partial}{\partial t}(tM div\, \boldsymbol{H}_o) = 0 \end{aligned}$$

since $\boldsymbol{E}_o, \boldsymbol{H}_o$ are solenoidal by assumption. This proves (4.37).

Remark 3. Theorem 4.3.4 shows that in order to guarantee that the solution is C^1 with respect to $\boldsymbol{x}$ for every $t > 0$ one must assume that $(\boldsymbol{E}', \boldsymbol{H}')$ is C^2 initially, for $t = 0$. It follows that, if we freeze the field at some time $t_o > 0$ and attempt to use $\boldsymbol{E}'(\boldsymbol{x}, t_o)$ and $\boldsymbol{H}'(\boldsymbol{x}, t_o)$ as new initial data, we have no guarantee that $\boldsymbol{E}'(\boldsymbol{x}, t_o)$ and $\boldsymbol{H}'(\boldsymbol{x}, t_o)$ are in $C^2(\mathbb{R}^3)$. This phenomenon of loss of derivatives, which is well-known in connection with hyperbolic equations in two or more space variables [18], is related to focussing (formation of caustics), which may occur in physics when there are two or more space dimensions. Examples show that this loss is indeed a real possibility [2], [18]. From the mathematical point of view, the phenomenon is due to the presence of the terms

$$\frac{\partial}{\partial t}(tM\boldsymbol{E}_o) \equiv M\boldsymbol{E}_o + t\frac{\partial}{\partial t} M\boldsymbol{E}_o \quad , \quad \frac{\partial}{\partial t}(tM\boldsymbol{H}_o) \equiv M\boldsymbol{H}_o + t\frac{\partial}{\partial t} M\boldsymbol{H}_o$$

that depend also on the gradient of the initial data. Indeed, if f is any component of the vectors $\boldsymbol{E}_o$ or $\boldsymbol{H}_o$, we have

$$\frac{\partial}{\partial t} Mf = \frac{t}{4\pi} \int_\Omega \frac{\partial}{\partial t} f(\boldsymbol{x} + c_o t \boldsymbol{\nu})\, d\Omega = \frac{c_o t}{4\pi} \int_\Omega \boldsymbol{\nu} \cdot grad\, f(\boldsymbol{x} + r\boldsymbol{\nu})\, d\Omega$$

Thus in order to have a C^2 solution for all t, one must "control" also the gradient of the initial data, and the continuity of derivatives is not in general persistent in time, as it happens for the wave equation.

Remark 4. (domain of dependence). Eq. (4.44) yields a representation of the electromagnetic field $(\boldsymbol{E}', \boldsymbol{H}')$ via spherical means of the initial data of the type

$$tM curl\, \boldsymbol{H}_o = \frac{1}{4\pi c_o^2 t} \int_{\Sigma_{c_o t}} curl_y \boldsymbol{H}_o(\boldsymbol{y})\, dS_y$$

(where $\Sigma_{c_o t}$ is the sphere $|\boldsymbol{y}-\boldsymbol{x}| = c_o t$) and their time derivative. Therefore $\boldsymbol{E}'(\boldsymbol{x}_o,t)$ and $\boldsymbol{H}'(\boldsymbol{x}_o,t)$ depend only on the values of the initial data $\boldsymbol{E}_o(\boldsymbol{x}), \boldsymbol{H}_o(\boldsymbol{x})$ taken over the spherical surface

$$\mathcal{D}(\boldsymbol{x}_o, t) := \{\boldsymbol{y} \in \mathbb{R}^3 : |\boldsymbol{y} - \boldsymbol{x}_o| = c_o t\} \tag{4.45}$$

called domain of dependence of $(\boldsymbol{x}_o, t)$. The Maxwell equations satisfy the Huygens's principle, which says that the domain of dependence consists only of boundary points. This principle implies the existence of a rear wave front as well as a leading wave front and guarantees that a sharply localized initial state remains sharply localized at all times [2,18].

Remark 5. (Domain of influence and determinacy). If the initial data $\boldsymbol{E}_o(\boldsymbol{x}), \boldsymbol{H}_o(\boldsymbol{x})$ are assigned only in the ball $B_o : |\boldsymbol{x}-\boldsymbol{x}_o| \leq R$ in $\mathbb{R}^3$, the field $(\boldsymbol{E}',\boldsymbol{H}')$ is uniquely determined in the domain of determinacy

$$\mathcal{B}(B_o) := \big\{\boldsymbol{x} \in \mathbb{R}^3,\, t > 0 : |\boldsymbol{x} - \boldsymbol{x}_o| \leq R - c_o t\big\} \tag{4.46}$$

(cfr. the proof of Theorem 4.3.1), and is influenced by these data in the larger domain of influence

$$\mathcal{J}(B_o) := \{\boldsymbol{x} \in \mathbb{R}^3,\, t > 0 : c_o t - R \leq |\boldsymbol{x} - \boldsymbol{x}_o| \leq c_o t + R\} \tag{4.47}$$

bounded by the leading wavefront $|\boldsymbol{x}-\boldsymbol{x}_o| = c_o t + R$ and, for $t > R/c_o$, by the rear wavefront $|\boldsymbol{x}-\boldsymbol{x}_o| = c_o t - R$. In the case at hand both wavefronts are spherical.

Remark 6. (Propagation of the support). If the initial data have compact support B, the solution has, for all $t > 0$, compact support $\mathcal{J}(B)$, contained between the leading and rear wavefronts. In other words, the "signals" propagate with finite speed c_o.

Remark 7. (Huygens' construction). More generally, if the support of the initial data is any compact set K in $\mathbb{R}^3$, the field (4.44) vanishes identically unless $\delta < t < \Delta$, where

$$\delta := \min_{\boldsymbol{y}\, \epsilon K} |\boldsymbol{x} - \boldsymbol{y}|/c_o \quad , \qquad \Delta := \max_{\boldsymbol{y}\, \epsilon K} |\boldsymbol{x} - \boldsymbol{y}|/c_o$$

are the times of passage through $\boldsymbol{x}$ of the leading and rear wave front, respectively. As first remarked by Huygens, the leading and rear wavefronts can be constructed at every time t by taking the envelopes of the family of spheres with centers $\boldsymbol{y} \in \partial K$ and radii $c_o t$ [18].

Remark 8. The treatment presented in this section can be immediately extended to any homogeneous dielectric with constant permittivity ϵ and is an adaptation of mathematical concepts and methods developed for the wave equation in 3D [2].

4.4 Radiation problem

We come now to the so-called radiation problem for $(\boldsymbol{E}'',\boldsymbol{H}'')$, in which the initial data are all zero and the electric current $\boldsymbol{J}=\boldsymbol{J}_r(\boldsymbol{x},t)$ is assigned for $t>0,\boldsymbol{x}\in\mathbb{R}^3$. Dropping all the superscripts for brevity, the electromagnetic field $\boldsymbol{E}=\boldsymbol{E}'',\boldsymbol{H}=\boldsymbol{H}''$ satisfies the inhomogeneous Maxwell equations

$$\mu_o\frac{\partial \boldsymbol{H}}{\partial t} = -curl\,\boldsymbol{E} \tag{4.48}$$

$$\epsilon_o\frac{\partial \boldsymbol{E}}{\partial t} = curl\,\boldsymbol{H} - \boldsymbol{J}_r(\boldsymbol{x},t) \tag{4.49}$$

$$div\,\boldsymbol{E} = \rho(\boldsymbol{x},t)/\epsilon_o \ , \ div\,\boldsymbol{H} = 0 \tag{4.50}$$

for $\boldsymbol{x}\in\mathbb{R}^3$, $t>0$, with the homogenous initial conditions

$$\boldsymbol{E}(\boldsymbol{x},0) = \boldsymbol{0} \ , \ \mathbf{H}(\boldsymbol{x},0) = \boldsymbol{0} \ , \ \rho(\boldsymbol{x},0) = 0 \quad (\boldsymbol{x}\in\mathbb{R}^3) \tag{4.51}$$

for $t=0$. By virtue of eq. (4.36), the charge density is determined by the given current according to the relation

$$\rho(\boldsymbol{x},t) = -\int_0^t div\,\boldsymbol{J}_r(\boldsymbol{x},\tau)d\tau \qquad (\boldsymbol{x}\in\mathbb{R}^3,\ t\geq 0). \tag{4.52}$$

4.4.1 Electrodynamic potentials and gauge transformation.

The radiation problem has at most one solution $(\boldsymbol{E},\boldsymbol{H})$, as proven in Theorem 4.3.1. In order to prove existence and find a representation formula of the solution, we will proceed formally to begin with and we will justify our passages later on. The radiation problem (4.48)-(4.51) can be conveniently reformulated if we introduce two potentials for the electromagnetic field: first of all, since $\boldsymbol{H}$ is solenoidal in $\mathbb{R}^3$ (eq. (4.51)), there exists a vector potential $\boldsymbol{V}'(\boldsymbol{x},t)$ for $\boldsymbol{B}=\mu_o\boldsymbol{H}$, such that

$$\mu_o\boldsymbol{H}(\boldsymbol{x},t) = curl\ \boldsymbol{V}'(\boldsymbol{x},t) \quad \mathbf{x}\in\mathbb{R}^3,\ t\geq 0 \tag{4.53}$$

Substituting eq. (4.53) into (4.48) we obtain the equation

$$curl\,(\boldsymbol{E} + \frac{\partial\boldsymbol{V}'}{\partial t}) = \mathbf{0} \tag{4.54}$$

for $\boldsymbol{x}\in\mathbb{R}^3$, $t>0$, which implies the existence of a scalar potential $u'(\boldsymbol{x},t)$ such that

$$\boldsymbol{E}(\boldsymbol{x},t) = -grad\,u' - \frac{\partial\,\boldsymbol{V}'}{\partial t} \tag{4.55}$$

for $\boldsymbol{x}\in\mathbb{R}^3$, $t>0$. On the other hand, the vector potential is clearly determined up to the gradient of an arbitrary (differentiable) scalar function $\phi(\boldsymbol{x},t)$: if $\boldsymbol{V}'$ is a vector potential, then

$$\boldsymbol{V} := \boldsymbol{V}' + grad\,\phi$$

is also a vector potential, in the sense that

$$\boldsymbol{H}(\boldsymbol{x},t) = \frac{1}{\mu_o}curl\ \boldsymbol{V}(\boldsymbol{x},t) \tag{4.56}$$

is unchanged. If we add this gradient term, eq. (4.55) shows that the electric field is then given by the equation

$$\boldsymbol{E}(\boldsymbol{x},t) = -grad\,u' - \frac{\partial\,\boldsymbol{V}}{\partial t} + grad\frac{\partial\,\phi}{\partial t} \equiv -grad\,(u' - \frac{\partial\,\phi}{\partial t}) - \frac{\partial\,\boldsymbol{V}}{\partial t}$$

which has the same form of eq. (4.55)

$$\boldsymbol{E}(\boldsymbol{x},t) = -grad\,u - \frac{\partial\boldsymbol{V}}{\partial t} \tag{4.57}$$

if we define the new scalar potential

$$u := u' - \frac{\partial\phi}{\partial t}$$

$\phi(\boldsymbol{x}, t)$ is called a gauge function. We have shown that if u', $\boldsymbol{V}'$ are two fixed potentials and $\phi(\boldsymbol{x}, t)$ is an arbitrary gauge function , the fields $\boldsymbol{E}$, $\boldsymbol{H}$ can also be represented by eqs. (4.56) and (4.57), provided the new potentials u, $\boldsymbol{V}$ are defined in terms of $u', \boldsymbol{V}'$ by the gauge transformation

$$u = u' - \frac{\partial \phi}{\partial t} \ , \quad \boldsymbol{V} = \boldsymbol{V}' + grad\, \phi \tag{4.58}$$

A convenient choice of the gauge function ϕ is the following. Differentiating the first equation (4.58) with respect to t and taking the divergence of the second after multiplying by $c_o^2 = (\epsilon_o \mu_o)^{-1}$ we obtain

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{\partial u'}{\partial t} - \frac{\partial u}{\partial t} \ , \quad c_o^2 \Delta_3 \phi = c_o^2 div\, \boldsymbol{V} - c_o^2 div\, \boldsymbol{V}'$$

For fixed u' and V$'$, we choose ϕ as a solution of the inhomogeneous wave equation

$$\frac{\partial^2 \phi}{\partial t^2} - c_o^2 \Delta_3 \phi = c_o^2 div\, \boldsymbol{V}' + \frac{\partial u'}{\partial t}$$

Then u and $\boldsymbol{V}$ satisfy the Lorentz condition

$$\frac{\partial u}{\partial t} + c_o^2 div\, \boldsymbol{V} = 0 \tag{4.59}$$

and substituting eqs. (4.56), (4.57) and (4.59) into eq. (4.49) yields

$$\boldsymbol{0} = \epsilon_o \frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{J}_r - curl\, \boldsymbol{H} = -\epsilon_o\, grad\, \frac{\partial u}{\partial t} - \epsilon_o \frac{\partial^2 \boldsymbol{V}}{\partial t^2} + \boldsymbol{J}_r - \mu_o^{-1} curl\, curl\, \boldsymbol{V}$$
$$= \epsilon_o c_o^2 grad\, div\, \boldsymbol{V} - \mu_o^{-1} curl\, curl\, \boldsymbol{V} - \epsilon_o \frac{\partial^2 \boldsymbol{V}}{\partial t^2} + \boldsymbol{J}_r$$

Multiplying by $\epsilon_o^{-1} = c_o^2 \mu_o$ we thus see that $\boldsymbol{V}$ satisfies (in cartesian coordinates) the nonhomogeneous vector wave equation

$$\frac{\partial^2 \boldsymbol{V}}{\partial t^2} - c_o^2 \Delta_3 \boldsymbol{V} = c_o^2 \mu_o \boldsymbol{J}_r(\boldsymbol{x}, t) \tag{4.60}$$

In this way eqs. (4.48), (4.49) and the second eq. (4.51) are satisfied, whereas the first eq. (4.50), eq. (57) and eq. (4.59) yield the nonhomogeneous scalar wave equation for u

$$\frac{\partial^2 u}{\partial t^2} - c_o^2 \Delta_3 u = \frac{c_o^2}{\epsilon_o} \rho(\boldsymbol{x}, t) \tag{4.61}$$

where $\rho(\boldsymbol{x},t)$ is given by eq. (4.52). The homogeneous initial conditions (4.51) for $t=0$ are satisfied if

$$\boldsymbol{V}(\boldsymbol{x},0) = \boldsymbol{V}_t(\boldsymbol{x},0) = 0 \;\; , \;\; u(\boldsymbol{x},0) = u_t(\boldsymbol{x},0) = 0 \tag{4.62}$$

and all we need is to find the potentials u, $\boldsymbol{V}$.

4.4.2 Retarded potentials.

The potentials u, $\boldsymbol{V}$ were first determined by Kirchhoff in closed form.

Theorem 4.4.1 *Let $\boldsymbol{J}_r(\boldsymbol{x},t)$ be a given function of class $C^2(\mathbb{R}^4)$ with compact support $K\subset\mathbb{R}^3$ for every $t\geq 0$, and let $\rho(\boldsymbol{x},t)$ be given by (4.52). Then the Cauchy problems (4.60), (4.61), (4.62) have unique C^2 solutions $u,\boldsymbol{V}$ which satisfy the Lorentz condition (4.59) and are given by Kirchhoff's retarded potentials*

$$u(\boldsymbol{x},t) = \frac{1}{4\pi\epsilon_o}\int_{|\boldsymbol{y}-\boldsymbol{x}|\leq c_o t} \rho(\boldsymbol{y},t-\frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o})\,\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|} \tag{4.63}$$

$$\boldsymbol{V}(\boldsymbol{x},t) = \frac{\mu_o}{4\pi}\int_{|\boldsymbol{y}-\boldsymbol{x}|\leq c_o t} \boldsymbol{J}_r(\boldsymbol{y},t-\frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o})\,\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|} \tag{4.64}$$

Proof. The potential u can be written in the form

$$u(\boldsymbol{x},t) = \frac{1}{4\pi\epsilon_o}\int_0^{c_o t}\frac{dr}{r}\int_{\Sigma_r}\rho(\boldsymbol{y},t-\frac{r}{c_o})\,dS_y \equiv \frac{1}{4\pi\epsilon_o}\int_o^t\frac{d\theta}{t-\theta}\int_{\Sigma_r}\rho(\boldsymbol{y},\theta)\,dS_y$$
$$\equiv \frac{c_o^2}{4\pi\epsilon_o}\int_o^t d\theta\,(t-\theta)\int_\Omega \rho(\boldsymbol{x}+c_o(t-\theta)\boldsymbol{\nu},\theta)\,d\Omega$$

where r, θ are integration variables related to t by $\theta = t-r/c_o$, so that

$$r = c_o(t-\theta) \quad , \quad \boldsymbol{y} = \boldsymbol{x}+r\boldsymbol{\nu} = \boldsymbol{x}+c_o(t-\theta)\boldsymbol{\nu} \;,\;\; dS_y = r^2 d\Omega = c_o^2(t-\theta)^2 d\Omega$$

and Σ_r is the sphere $|\boldsymbol{y}-\boldsymbol{x}| = r\equiv c_o(t-\theta)$. Under the stated assumptions we have $u(\boldsymbol{x},0)=0$ and

$$\Delta_3 u(\boldsymbol{x},t) = \frac{c_o^2}{4\pi\epsilon_o}\int_o^t d\tau\,(t-\theta)\int_\Omega \Delta_3\rho(\boldsymbol{y},\theta)\,d\Omega \equiv \frac{1}{4\pi\epsilon_o}\int_o^t\frac{d\theta}{t-\theta}\int_{\Sigma_r}\Delta_3\rho(\boldsymbol{y},\theta)\;dS_y$$

Moreover,for $t > 0$

$$u_t = \frac{c_o^2}{4\pi\epsilon_o}\int_o^t d\theta \int_\Omega \rho(\boldsymbol{x} + c_o(t-\theta)\boldsymbol{\nu}, \theta)\, d\Omega + \frac{c_o^3}{4\pi\epsilon_o}\int_o^t d\theta(t-\theta)\int_\Omega \boldsymbol{\nu}\boldsymbol{\cdot} grad\, \rho(\boldsymbol{y}, \theta)\, d\Omega$$
$$= \frac{c_o^2}{4\pi\epsilon_o}\int_o^t d\theta \int_\Omega \rho(\boldsymbol{y}, \theta)\, d\Omega + \frac{c_o}{4\pi\epsilon_o}\int_o^t \frac{d\theta}{t-\theta}\int_{B_r} \Delta_3\rho(\boldsymbol{y}, \theta)\, d\boldsymbol{y}$$

where we have applied the divergence theorem to the ball $B_r : |\boldsymbol{y}-\boldsymbol{x}| \leq r \equiv c_o(t-\theta)$, whose boundary is the sphere Σ_r. Thus we have also $u_t(\boldsymbol{x}, 0) = 0$.

Since $d\boldsymbol{y}= drdS_y \;= c_o^2(t-\theta)^2 d\Omega d\theta$, it follows that

$$u_{tt} = \frac{c_o^2}{4\pi\epsilon_o}\rho(\boldsymbol{x}, t)\int_\Omega d\Omega + \frac{c_o^3}{4\pi\epsilon_o}\int_o^t d\theta \int_\Omega \boldsymbol{\nu}\boldsymbol{\cdot} grad\, \rho(\boldsymbol{y}, \theta)\, d\Omega$$
$$+ \frac{c_o^2}{4\pi\epsilon_o}\int_o^t \frac{d\theta}{t-\theta}\int_{\Sigma_r} \Delta_3\rho(\boldsymbol{y}, \theta)\, dS_y - \frac{c_o}{4\pi\epsilon_o}\int_o^t \frac{d\theta}{(t-\theta)^2}\int_{B_r} \Delta_3\rho(\boldsymbol{y}, \theta)\, d\boldsymbol{y}$$
$$= \frac{c_o^2}{\epsilon_o}\rho(\boldsymbol{x}, t) + c_o^2\Delta_3 u$$

and so u satisfies the wave equation (4.61). The proof for the vector potential $\boldsymbol{V}$ is entirely similar. The Lorentz condition (4.59) is also satisfied (Exercise 5). Thus all the formal passages carried out in eqs. (4.53)−(4.61) are justified, and the theorem is proven.

Eqs. (4.63), (4.64) and (4.52) show that $\rho \equiv 0$ implies $u \equiv 0$ and $\boldsymbol{J}_r \equiv \boldsymbol{0}$ implies $u \equiv$ V$\equiv\ 0$.

Corollary 4.4.2 *The unique C^1 solution of the radiation problem, given by eqs. (4.56), (4.57), (4.63) and (4.64), has compact support in $\mathbb{R}^3$ for all $t > 0$.*

Proof. Since $\boldsymbol{J}_r$ has compact support K in $\mathbb{R}^3$ for all t, the same is true for ρ, given by eq. (4.52). Eqs. (4.63) and (4.64) show then that, for all $t > 0$, $\boldsymbol{E}(\boldsymbol{x}, t)$ and $\mathbf{H}(\boldsymbol{x}, t)$ (as well as u and $\boldsymbol{V}$) have compact support

$$K_t = \{\boldsymbol{x} \in \mathbb{R}^3 : \inf_{\boldsymbol{y}\in K} |\boldsymbol{y} - \boldsymbol{x}| \leq c_o t\ \}$$

which propagates with speed c_o .

Corollary 4.4.3 *In the limit case of a linear antenna* Γ *driven by a current* $I(t)$ *we have*

$$u(\boldsymbol{x},t) \equiv 0 \quad , \quad \boldsymbol{V}(\boldsymbol{x},t) = \frac{\mu_o}{4\pi}\int_{\Gamma_t} I(t - \frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o})\, \boldsymbol{t}(\boldsymbol{y}) \frac{ds_y}{|\boldsymbol{y}-\boldsymbol{x}|}$$

where $\Gamma_t := \left\{\boldsymbol{y} \in \Gamma : |\boldsymbol{y}-\boldsymbol{x}| \leq c_o t\right\}$, *andthe electromagnetic field is given by*

$$\boldsymbol{E}(\boldsymbol{x},t) = -\frac{\partial \boldsymbol{V}}{\partial t} \qquad , \qquad \boldsymbol{H}(\boldsymbol{x},t) = \frac{1}{\mu_o} curl\ \boldsymbol{V}$$

Proof. In this case $\boldsymbol{J}_r(\boldsymbol{x},t) = I(t)\delta_\Gamma(\boldsymbol{x})\mathrm{t}(\boldsymbol{x})$ (cfr. (1.11)) and $\rho \equiv 0$, so that $u \equiv 0$. The assertion then follows from eqs. (4.56) and (4.57).

Theorem 4.4.1 has two further corollaries which show that transients are finite: namely, if the current $\boldsymbol{J}_r(\boldsymbol{x},t)$ is periodic in t and has compact support K in $\mathbb{R}^3$, the electromagnetic field is also periodic after a finite time

$$T(\boldsymbol{x}) := c_o^{-1} \sup_{\boldsymbol{y}\in\partial K} |\boldsymbol{y}-\boldsymbol{x}| \tag{4.65}$$

which depends on the distance of the observation point $\boldsymbol{x}$ from the boundary of K. In particular, $T(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in K$.

Corollary 4.4.4 *Suppose* $\boldsymbol{J}_r$ *has the form*

$$\boldsymbol{J}_r(\boldsymbol{x},t) = \boldsymbol{J}_o(\boldsymbol{x})e^{i\omega t}$$

where $\boldsymbol{J}_o(\boldsymbol{x})$ *has compact support* K. *For* $t \geq T(\boldsymbol{x})$ *we have*

$$\boldsymbol{V}(\boldsymbol{x},t) = \boldsymbol{V}_o(\boldsymbol{x})e^{i\omega t} \;, \qquad \boldsymbol{H}(\boldsymbol{x},t) = \frac{1}{\mu_o} curl\ \boldsymbol{V}_o(\boldsymbol{x})\, e^{i\omega t} \tag{4.66}$$

where $\boldsymbol{V}_o(\boldsymbol{x})$ *is a suitable function of* $\boldsymbol{x}$, *and if* $\boldsymbol{J}_o$ *is solenoidal*

$$u(\boldsymbol{x},t) \equiv 0 \;, \qquad \boldsymbol{E}(\boldsymbol{x},t) = -i\omega\, \boldsymbol{V}_o(\boldsymbol{x})e^{i\omega t} \tag{4.67}$$

Proof. For $t \geq T(\boldsymbol{x})$ the integrals in eqs. (4.63) and (4.64) are carried over all of K so that

$$\boldsymbol{V}(\boldsymbol{x},t) = \frac{\mu_o}{4\pi}\int_K \boldsymbol{J}_o(\boldsymbol{y})e^{i\omega(t-\frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o})}\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|} = e^{i\omega t}\frac{\mu_o}{4\pi}\int_K \boldsymbol{J}_o(\boldsymbol{y})e^{-i\omega\frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o}}\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|}$$

and so

$$\boldsymbol{V}_o(\boldsymbol{x}) := \frac{\mu_o}{4\pi}\int_K \boldsymbol{J}_o(\boldsymbol{y})e^{-i\omega\frac{|\boldsymbol{y}-\boldsymbol{x}|}{c_o}}\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|} \tag{4.68}$$

Eq. (4.52) yields

$$\rho(\boldsymbol{x},t) = -\int_0^t div\boldsymbol{J}_r(\boldsymbol{x},\tau)d\tau \equiv -div\,\boldsymbol{J}_o(\boldsymbol{x})\int_0^t e^{i\omega\tau}d\tau \tag{4.69}$$

$$\equiv -\frac{i}{\omega}div\,\boldsymbol{J}_o(\boldsymbol{x})(e^{i\omega t}-1)$$

If $div\boldsymbol{J}_o = 0$ we have $\rho \equiv 0$, so that $u(\boldsymbol{x},t) \equiv 0$ and eqs. (4.56), (4.57) reduce to (4.66), (4.67).

Corollary 4.4.5 *Suppose* $\boldsymbol{J}_r(\boldsymbol{x},t) = \boldsymbol{J}_o(\boldsymbol{x})$ *is time-independent, solenoidal and with compact support. Then*

(i) $\boldsymbol{H}(\boldsymbol{x},t) = \boldsymbol{H}(\boldsymbol{x})$ *for* $t \geq T(\boldsymbol{x})$, *where* $\boldsymbol{H}(x)$ *is given by the Biot-Savart integral*

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{1}{4\pi}curl\int_K \boldsymbol{J}_o(\boldsymbol{y})\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|} \tag{4.70}$$

(ii) $\boldsymbol{E}(\boldsymbol{x},t) \equiv \boldsymbol{0}$ *for* $t \geq T(\boldsymbol{x})$, *in particular for all* $\boldsymbol{x} \in K$, $t \geq 0$.

Proof. Proceeding as in eq. (4.69) we see that $\rho \equiv 0$, so that $u \equiv 0$. For $t \geq T(\boldsymbol{x})$ eq. (4.64) shows that $\boldsymbol{V}(\boldsymbol{x},t) = \boldsymbol{V}(\boldsymbol{x})$ is time-independent, with

$$\boldsymbol{V}(\boldsymbol{x}) = \frac{\mu_o}{4\pi}\int_K \boldsymbol{J}_o(\boldsymbol{y})\frac{\boldsymbol{dy}}{|\boldsymbol{y}-\boldsymbol{x}|}$$

Eqs. (4.56) and (4.57) yield then $\boldsymbol{E}(\boldsymbol{x},t) \equiv \boldsymbol{0}$ and $\boldsymbol{H}(\boldsymbol{x},t) = \boldsymbol{H}(\boldsymbol{x})$ for $t \geq T(\boldsymbol{x})$, where $\boldsymbol{H}(\boldsymbol{x})$ is given by (4.70) and coincides with the Biot-Savart magnetic field (3.23).

Remark 9. If $\boldsymbol{J}_r$ (and hence ρ) are assumed identically zero for $t \leq 0$, Kirchhoff's retarded potentials (4.63) and (4.64) take the form of time-dependent volume potentials

$$u(\boldsymbol{x},t) = \frac{1}{4\pi\epsilon_o}\int_{\mathbb{R}^3}\rho(\boldsymbol{y},t-\frac{|\boldsymbol{x}-\boldsymbol{y}|}{c_o})\frac{\boldsymbol{dy}}{|\boldsymbol{x}-\boldsymbol{y}|}$$

$$\boldsymbol{V}(\boldsymbol{x},t) = \frac{\mu_o}{4\pi}\int_{\mathbb{R}^3} \boldsymbol{J}_r(\boldsymbol{y}, t - \frac{|\boldsymbol{x}-\boldsymbol{y}|}{c_o})\frac{\boldsymbol{dy}}{|\boldsymbol{x}-\boldsymbol{y}|}$$

(see §2.1) with a time delay depending on $\boldsymbol{y}$. This explains the name of retarded potentials and shows that the influence of the current $\boldsymbol{J}_r$ at the point $\boldsymbol{y}$ travels with finite speed c_o, as in Remark 5 above.

Remark 10. Replacing ϵ_o with ϵ, μ_o with μ and c_o with c in the previous formulae, we immediately obtain the solution of the radiation problem for a homogeneous dielectric occupying all the space $\mathbb{R}^3$. The Cauchy problem for a conductor ($\gamma > 0$) will be treated in §4.6.

Remark 11. Corollary 4.4.5 (*ii*) is in contradiction with the results of Chapter 3. This is due to the fact that $\boldsymbol{E}$ is assumed zero initially and must be stationary inside K, as $\boldsymbol{J}_r(\boldsymbol{x},t) = \boldsymbol{J}_o(\boldsymbol{x})$ is stationary. Hence $\boldsymbol{E} \equiv \boldsymbol{0}$ for all $\boldsymbol{x} \in K$ and all $t \geq 0$. Outside K the electric field $\boldsymbol{E}(\boldsymbol{x},t)$ is zero as soon as the magnetic field $\boldsymbol{H}(\boldsymbol{x},t)$ becomes stationary, i.e. for $t \geq T(\boldsymbol{x})$.

Note also that the conductor K need not be toroidal here.

Remark 12. The electromagnetic field can also be represented by means of a single vector potential, called Hertz vector, or by means of two scalar potentials [32, 43]. The Hertz vector is especially appropriate to deal with radiation from linear antennas.

4.5 Telegraph equation

A transmission line, consisting of a pair of very long parallel plates at distance d from one another and carrying opposite currents, plays an important role in the transmission of electric signals, as a simple model of the telegraph [6]. Even if condition (1.81) is violated in application to telegraphy, because of the exceeding length of the line, one can still apply the quasi-stationary approximation locally by extending the equations of electric circuits to the case of distributed parameters.

Suppose that the two rectangular plates are parallel to the $x-$axis, and that their length l, width $\mathcal{L}$ and distance d satisfy $l \gg \mathcal{L} \gg d$. Two opposing points at the same abscissa x have equal and opposite currents $\pm I(x,t)$ and surface charge densities $\pm\sigma(x,t) = \pm\chi(x,t)/\mathcal{L}$, where $\chi(x,t)$ is the charge

[6] in a more realistic model the two plates are replaced by a coaxial cable [43]

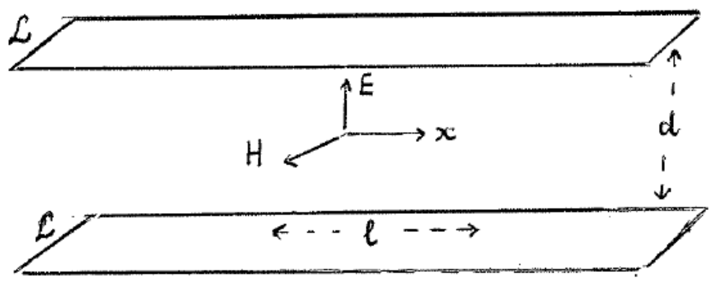

Figure 4.1: Model of a telegraphic line

density per unit length. The difference of potential $V(x,t)$ between the two points satisfies

$$\chi = \mathbb{C}'V$$

where $\mathbb{C}'$ is the capacity per unit length of the line. Supposing that the interposing medium is slightly conducting, this potential difference gives rise to a cross current $G'V(x,t)$,where G' is the "conductance" of the line, a parameter that can be partly controlled by augmenting or diminishing the insulation between the two plates. This cross current contributes to modify the linear charge density $\chi(x,t)$. In this way the transmission line can be viewed as an imperfectly insulated condenser. The continuity equation for a small (infinitesimal) stretch, centered at the abscissa x, of the upper plate is

$$\frac{\partial \chi}{\partial t} + G'V = -\frac{\partial I}{\partial x}$$

or

$$\mathbb{C}'\frac{\partial V}{\partial t} + \frac{\partial I}{\partial x} + G'V = 0 \tag{4.71}$$

Note that for $G' = 0$ and χ independent of time this equation reduces to $\frac{\partial I}{\partial x} = 0$ and implies $I = I(t)$, as is the case of the electric circuits with concentrated parameters studied in §3.7. The balance equation for the voltage drops reads

$$L'\frac{\partial I}{\partial t} = -\frac{1}{\mathbb{C}'}\frac{\partial \sigma}{\partial x} - \mathcal{R}'I$$

or

$$L'\frac{\partial I}{\partial t} + \frac{\partial V}{\partial x} + \mathcal{R}'I = 0 \tag{4.72}$$

where $\mathcal{R}'$ is the resistance per unit length and L' the inductance per unit length of the transmission line (Exercise 6). The parameters $\mathbb{C}'$, L', $\mathcal{R}'$ can also be controlled by varying the geometry.

The partial differential equations (4.71), (4.72) form a linear first order hyperbolic system in the two unknowns $I(x,t)$ and $V(x,t)$. Substituting the expression for V obtained from eq. (4.71)

$$V = -G'^{-1}\,\frac{\partial I}{\partial x} \;\; - G'^{-1}\mathbb{C}'\frac{\partial V}{\partial t}$$

into eq. (4.72) yields

$$L'\,\frac{\partial I}{\partial t} + \mathcal{R}' I - \frac{1}{G'}\,\frac{\partial^2 I}{\partial x^2} \;\; - \frac{\mathbb{C}'}{G'}\frac{\partial^2 V}{\partial t \partial x} \; = 0$$

where,by differentiating eq. (4.72) with respect to t, the last term can be written in terms of I as

$$\frac{\partial^2 V}{\partial t \partial x} = -L'\,\frac{\partial^2 I}{\partial t^2} \;\; - \mathcal{R}'\frac{\partial I}{\partial t}$$

In this way we arrive at the damped wave equation for $I(x,t)$

$$\frac{\partial^2 I}{\partial t^2} + (\alpha+\beta)\,\frac{\partial I}{\partial t} + \alpha\beta\, I = a^2\,\frac{\partial^2 I}{\partial x^2} \tag{4.73}$$

where

$$\alpha := \frac{G'}{\mathbb{C}'} \quad , \qquad \beta := \frac{\mathcal{R}'}{L'} \;\; , \quad a := \frac{1}{\sqrt{\mathbb{C}' L'}}$$

The parameters $\mathcal{R}'$ and G' can be adjusted so that

$$\frac{\mathcal{R}'}{L'} = \frac{G'}{\mathbb{C}'} \tag{4.74}$$

With this choice $\alpha = \beta$ and eq. (4.73) becomes the telegraph equation

$$\frac{\partial^2 I}{\partial t^2} + 2\beta\,\frac{\partial I}{\partial t} + \beta^2 I = a^2\,\frac{\partial^2 I}{\partial x^2} \tag{4.75}$$

first derived by Heaviside in connection with the propagation of telegraphic signals in underwater cables. The crucial importance of taking $\alpha = \beta$ lies in the fact that the telegraph equation admits undistorted traveling wave solution of arbitrary form $w(x)$, damped in time

$$I(x,t) = e^{-\beta t} w(x - at) \tag{4.76}$$

or equivalently in space

$$I(x,t) = e^{-\beta x/a} w(x - at) \tag{4.77}$$

(Exercise 7 and 8). It follows that the telegraphic signal is transmitted with speed a along the line, attenuated but undistorted, and is received, weakened but fully recognizable, at the end of the line. (In contrast, if $\alpha \neq \beta$ no undistorted telegraph signal would be possible.)

The corresponding traveling wave solution for V is obtained from eqs. (4.71), (4.72) as

$$V(x,t) = \sqrt{\frac{L'}{\mathbb{C}'}} I(x,t) \tag{4.78}$$

(Exercise 9), where $\sqrt{L'/\mathbb{C}'}$ is the wave resistance. This implies a remarkable practical consequence: if the transmission line is closed at one end by means of an ohmic resistance $\mathcal{R}_c = \sqrt{L'/\mathbb{C}'}$, the current and the voltage remain continuous and no reflected signal arises.

What is the value of the propagation speed a of the telegraphic signal in this model? To answer this question we remark that, since $l \gg \mathcal{L} \gg d$ by assumption, the electric field $\boldsymbol{E}$ will be approximately orthogonal to the plates and the magnetic field $\boldsymbol{H}$ orthogonal to both $\boldsymbol{E}$ and the $x-$axis (see Fig. 4.1). By force of eqs. (1.3) and (1.4), eq. (2.76) and eqs. (3.33) and (3.82) we may write

$$V = |\boldsymbol{E}|d \ , \quad \chi \cong \epsilon|\boldsymbol{E}|\mathcal{L} \, , \quad I \cong |\boldsymbol{H}|\mathcal{L} \ , \quad \mathbb{C}' \simeq \frac{\chi}{V} \ , \quad \mu|\boldsymbol{H}|d \cong L'I$$

It follows that

$$\mathbb{C}' \cong \frac{\chi}{V} \cong \frac{\epsilon|\boldsymbol{E}|\mathcal{L}}{|\boldsymbol{E}|d} = \frac{\epsilon\mathcal{L}}{d} \, , \qquad L' \cong \frac{\mu|\boldsymbol{H}|d}{I} \cong \mu\frac{d}{\mathcal{L}}$$

and therefore

$$a = \frac{1}{\sqrt{\mathbb{C}'L'}} \cong \frac{1}{\sqrt{\epsilon\mu}} = c$$

Thus, within the limit of our approximations, the telegraphic signal travels with the speed of light.

4.6 Weak solutions of the Cauchy problem

We resume the study of the Cauchy problem for the inhomogeneous Maxwell equations in a homogeneous medium with physical constants ϵ, $\mu, \gamma \geq 0$

$$\epsilon \frac{\partial \boldsymbol{E}}{\partial t} - curl\,\boldsymbol{H} + \gamma \boldsymbol{E} = \boldsymbol{F}(\boldsymbol{x}, t) \tag{4.79}$$

$$\mu \frac{\partial \boldsymbol{H}}{\partial t} + curl\,\boldsymbol{E} = \boldsymbol{G}(\boldsymbol{x}, t) \quad (t > 0,\ \boldsymbol{x} \in \mathbb{R}^3)$$

with source terms of general form $\boldsymbol{F}$, $\boldsymbol{G}$ assigned for $t > 0$, $\boldsymbol{x} \in \mathbb{R}^3$, and initial data

$$\boldsymbol{E}(\boldsymbol{x}, 0) = \boldsymbol{E}_o(\boldsymbol{x}) \ , \quad \boldsymbol{H}(\boldsymbol{x}, 0) = \boldsymbol{H}_o(\boldsymbol{x}) \tag{4.80}$$

assigned for $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$. Note that we are including the case of conductors ($\gamma > 0$).

Proposition 4.6.1 *The Cauchy problem (4.79)−(4.80) has at most one solution* $(\boldsymbol{E},\boldsymbol{H}) \in C^o(\mathbb{R}^3 \times [0, +\infty)) \cap C^1(\mathbb{R}^3 \times (0, +\infty))$.

Proof. Proceeding exactly as in the proof of Theorem 4.3.1 we arrive at the inequality

$$\frac{d\mathcal{E}[\mathcal{B}(t)]}{dt} \leq -\frac{1}{2} \int_{\partial \mathcal{B}(t)} \left(\sqrt[4]{\frac{\epsilon_o}{\mu_o}} |\boldsymbol{E}| - \sqrt[4]{\frac{\mu_o}{\epsilon_o}} |\boldsymbol{H}| \right)^2 dS - \int_{\mathcal{B}(t)} \gamma |\boldsymbol{E}|^2 d\boldsymbol{x} \ \leq \ 0$$

and the same conclusions follow.

Let

$$x_o := ct\,, \quad \mathrm{X} = (x_o, \boldsymbol{x}), \ \boldsymbol{w} = \boldsymbol{w}(\boldsymbol{X}) = (w_1, ..., w_6) := (\sqrt{\epsilon} \boldsymbol{F}(\boldsymbol{x}, t), \sqrt{\mu} \boldsymbol{G}(\boldsymbol{x}, t))$$
$$\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{X}) = (u_1, ..., u_6) := (\sqrt{\epsilon} \boldsymbol{E}(\boldsymbol{x}, t), \sqrt{\mu} \boldsymbol{H}(\boldsymbol{x}, t)) \quad , \quad \boldsymbol{u}_o(\boldsymbol{x}) := (\sqrt{\epsilon} \boldsymbol{E}_{\mathrm{o}}, \sqrt{\mu} \boldsymbol{H}_o)$$

where c is defined by eq. (4.5). The initial conditions (4.80) become then

$$\boldsymbol{u}(0, \boldsymbol{x}) = \boldsymbol{u}_o(\boldsymbol{x}) \qquad (\mathrm{x} \in \mathbb{R}^3) \tag{4.81}$$

and (4.79) takes the form of a linear system of first order partial differential equations with constant coefficients [7]

$$Lu := u_{x_o} + \sum_{k=1}^{3} \mathbb{A}_k u_{x_k} + \mathbb{B}u = w \qquad (x_o > 0,\ x \in \mathbb{R}^3) \tag{4.82}$$

where $\mathbb{A}_k$, $\mathbb{B}$ are constant 6×6 real matrices defined by

$$\mathbb{A}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbb{A}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.83}$$

$$\mathbb{A}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} \gamma' & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma' & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma' & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.84}$$

with

$$\gamma' := \gamma\sqrt{\mu/\epsilon} \geq 0$$

Since the matrices $\mathbb{A}_k$ are symmetric, and the system has normal form with respect to x_o, the linear differential operator L is said to be symmetric hyperbolic with respect to the variable $x_o = ct$ [18]. This symmetry property has important consequences.

4.6.1 Symmetric hyperbolic systems: energy estimates.

We begin by deriving an energy estimate in this context, following the pattern of §1.5. For $T > 0$, let

$$R_T := [0, cT) \times \mathbb{R}^3$$

[7] $u_{x_k} = \partial u/\partial x_k$, a.s.o.

We denote by

$$\|\boldsymbol{u}(x_o,\cdot)\| := \sqrt{\int_{\mathbb{R}^3} |\boldsymbol{u}|^2\, d\boldsymbol{x}}$$

the norm of $\boldsymbol{u}(x_o,\boldsymbol{x})$ in $L^2(\mathbb{R}^3)$, and by

$$\|\boldsymbol{w}\|_{T} := \sqrt{\int_0^{cT} \|\boldsymbol{w}(x_o,\cdot)\|^2\, dx_o} \equiv \sqrt{\int_0^{cT}\int_{\mathbb{R}^3} |\boldsymbol{w}(x_o,\boldsymbol{x})|^2\, dx_o d\boldsymbol{x}} \tag{4.85}$$

the norm in $L^2(R_T)$. The energy $\mathcal{E}(x_o)$ for the entire space $\mathbb{R}^3$ associated to a solution $\boldsymbol{u}(x_o,\boldsymbol{x})$ of (4.81), (4.82) is the squared norm $\|\boldsymbol{u}(x_o,\cdot)\|^2$.

Proposition 4.6.2 *Suppose* $\boldsymbol{u}(x_o,\boldsymbol{x})$ *is a solution of (4.81), (4.82) of class* $C^o([0,T)\times\mathbb{R}^3)\cap C^1((0,T)\times\mathbb{R}^3)$ *with compact support in* $\mathbb{R}^3$ *for every* $x_o = ct \in [0,T]$*, where* $T>0$ *is arbitrary. Then the energy estimate holds*

$$\|\boldsymbol{u}(cT,\cdot)\|^2 \le e^{\,McT}(\|\boldsymbol{u}_o\|^2 + M\|L\boldsymbol{u}\|_{T}^2) \tag{4.86}$$

where $\boldsymbol{u}_o = \boldsymbol{u}(0,\boldsymbol{x})$,

$$M := \begin{cases} 1 & if \quad \|L\boldsymbol{u}\| > 0 \\ 0 & if \quad \|L\boldsymbol{u}\| = 0 \end{cases} \tag{4.87}$$

and $L\boldsymbol{u}=\boldsymbol{w}$.

Proof. Under our assumptions $\boldsymbol{u}$ and $L\boldsymbol{u}$ are in $L^2(R_T)$. The symmetry of the constant matrices $\mathbb{A}_k$ implies the identity

$$\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u}_{x_k} \equiv (\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u})_{x_k} - \boldsymbol{u}_{x_k}\cdot\mathbb{A}_k\boldsymbol{u} = (\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u})_{x_k} - \mathbb{A}_k\boldsymbol{u}_{x_k}\cdot\boldsymbol{u}$$

that is to say

$$\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u}_{x_k} \equiv \frac{1}{2}(\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u})_{x_k}$$

Multiplying eq. (4.82) scalarly by $\boldsymbol{u}$ yields then for $x_o > 0$

$$\boldsymbol{u}\cdot\boldsymbol{u}_{x_o} + \frac{1}{2}\sum_{k=1}^{3} (\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u})_{x_k} + \boldsymbol{u}\cdot\mathbb{B}\boldsymbol{u} = \boldsymbol{u}\cdot\boldsymbol{w}$$

and integrating over $\mathbb{R}^3$

$$\int_{\mathbb{R}^3} \boldsymbol{u}\cdot\boldsymbol{u}_{x_o} d\boldsymbol{x} + \frac{1}{2}\sum_{k=1}^{3}\int_{\mathbb{R}^3}\frac{\partial}{\partial x_k}(\boldsymbol{u}\cdot\mathbb{A}_k\boldsymbol{u})\,d\boldsymbol{x} \;+\int_{\mathbb{R}^3}\boldsymbol{u}\cdot\mathbb{B}\boldsymbol{u}\,d\mathbf{x} = \int_{\mathbb{R}^3}\boldsymbol{u}\cdot\boldsymbol{w}\,d\boldsymbol{x}$$

where $\boldsymbol{u}\cdot\boldsymbol{u}_{x_o} \equiv \frac{1}{2}\partial|\boldsymbol{u}|^2/\partial x_o$ and the second integral on the left hand side vanishes under our assumptions as a consequence of the Gauss Lemma. We thus obtain

$$\frac{1}{2}\frac{d}{dx_o}\int_{\mathbb{R}^3}|\boldsymbol{u}|^2 d\boldsymbol{x} = -\int_{\mathbb{R}^3}\boldsymbol{u}\cdot\mathbb{B}\boldsymbol{u}\,d\boldsymbol{x} + \int_{\mathbb{R}^3}\boldsymbol{u}\cdot\boldsymbol{w}\,d\boldsymbol{x}$$

The Schwartz inequality [39] implies that

$$2\int_{\mathbb{R}^3}\boldsymbol{w}\cdot\boldsymbol{u}\,d\boldsymbol{x} \le \;\|\boldsymbol{u}\|^2 + \|\boldsymbol{w}\|^2$$

and, since $-\boldsymbol{u}\cdot\mathbb{B}\boldsymbol{u}\le 0$ ($\mathbb{B}$ is positive semidefinite), we have

$$\frac{d}{dx_o}\|\boldsymbol{u}\|^2 \le M\|\boldsymbol{u}\|^2 + M\|\boldsymbol{w}\|^2$$

where the constant M is defined by eq. (4.87). It follows that

$$\frac{d}{dx_o}(e^{-Mx_o}\|\boldsymbol{u}\|^2) \le Me^{-Mx_o}\|\boldsymbol{w}\|^2 \le M\|\boldsymbol{w}\|^2$$

and integrating over x_o from 0 to $c\tau$ yields the equation

$$\|\boldsymbol{u}(c\tau,\cdot)\|^2 \le e^{Mc\tau}(\|\boldsymbol{u}_o\|^2 + M\|\boldsymbol{w}\|^2_{\tau}) \tag{4.88}$$

which for $\tau = T$ coincides with (4.86), since $\boldsymbol{w}= L\boldsymbol{u}$. This completes the proof.

Because $\|\boldsymbol{w}\|^2_{\tau} \le \|\boldsymbol{w}\|^2_{T}$ for $\tau \le T$, a further integration over $c\tau$ from 0 to cT in eq. (4.88) yields the estimate for the $L^2(R_T)-$norm of $\boldsymbol{u}$

$$\|\boldsymbol{u}\|^2_T \le \frac{1}{M}(e^{McT} - 1)\,(\|\boldsymbol{u}_o\|^2 + M\|\boldsymbol{w}\|^2_{T}) \tag{4.89}$$

if $M \neq 0$. If $\|\boldsymbol{w}\| = 0$ we have $M = 0$ and the energy estimate (4.86) reduces to the inequality

$$\|\boldsymbol{u}(cT,\cdot)\|^2 \le \|\boldsymbol{u}_o\|^2$$

or $\mathcal{E}(T) \leq \mathcal{E}(0)$. Similarly, if $\|\boldsymbol{w}\| = 0$ the inequality (4.89) becomes

$$\|\boldsymbol{u}\|_T^2 \leq \|\boldsymbol{u}_o\|_T^2 \equiv cT\|\boldsymbol{u}_o\|^2 \tag{4.90}$$

and for $\|\boldsymbol{u}_o\| = 0$ we have $\|\boldsymbol{u}\|_T = 0$, which implies (under more stringent assumptions than in Proposition 4.6.1) that a smooth solution of (4.81), (4.82) is unique.

4.6.2 Weak solution: existence and uniqueness.

The energy estimate implies existence and uniqueness, albeit with a suitable generalization of the concept of solution. We suppose from now on that the initial data $\boldsymbol{u}_o$ are zero:

$$\boldsymbol{u}(0, \boldsymbol{x}) \equiv \mathbf{0} \quad (\mathbf{x} \in \mathbb{R}^3) \tag{4.91}$$

This is not restrictive, since the initial data can be transformed into source terms by replacing $\boldsymbol{u}-\boldsymbol{u}_o$ with $\boldsymbol{u}$. In this way, excluding the trivial case, the constant M will be positive, and, recalling that $\boldsymbol{w}= L\boldsymbol{u}$, eq. (4.89) will take the form

$$\|\boldsymbol{u}\|_T^2 \leq K_T\|L\boldsymbol{u}\|_T^2$$

where $K_T := (e^{McT} - 1) > 0$. We denote $C_o^1(R_T)$ the subset of $C^1(R_T) \cap L^2(R_T)$ consisting of C^1 vector functions with compact support in R_T : such functions have compact support in $\mathbb{R}^3$ for all $x_o \in [0, cT)$ and vanish for $x_o = cT$ but not necessarily for $x_o = 0$. Moreover, we denote

$$((\boldsymbol{v}, \boldsymbol{w})) := \int_{R_T} \boldsymbol{v} \cdot \boldsymbol{w} \, d\boldsymbol{X}$$

the scalar product in $L^2(R_T)$, so that $((\boldsymbol{v},\boldsymbol{v})) = \|\boldsymbol{v}\|_T^2$. Scalarly multiplying system (4.82) by any test vector $\boldsymbol{v}\in C_o^1(R_T)$, applying the Gauss Lemma, and taking into account eq. (4.91) and the symmetry of the matrices $\mathbb{A}_k$ and

$\mathbb{B}$ yields

$$(4.92)\quad ((\boldsymbol{v},\boldsymbol{w})) = ((\boldsymbol{v},L\boldsymbol{u})) \;=\; ((\boldsymbol{v},\boldsymbol{u}_{x_o})) + \sum_{k=1}^{3}((\boldsymbol{v},\mathbb{A}_k\boldsymbol{u}_{x_k})) + ((\boldsymbol{v},\mathbb{B}\boldsymbol{u}))$$

$$= ((\boldsymbol{v},\boldsymbol{u}_{x_o})) + \sum_{k=1}^{3}((\mathbb{A}_k\boldsymbol{v},\boldsymbol{u}_{x_k})) + ((\mathbb{B}\boldsymbol{v},\boldsymbol{u}))$$

$$= -((\boldsymbol{v}_{x_o},\boldsymbol{u})) - \sum_{k=1}^{3}((\mathbb{A}_k\boldsymbol{v}_{x_k},\boldsymbol{u})) + ((\mathbb{B}\boldsymbol{v},\boldsymbol{u})) \;=((L^*\boldsymbol{v},\boldsymbol{u}))$$

where L^* is the adjoint operator of L, defined by

$$L^*\boldsymbol{v} := -\boldsymbol{v}_{x_o} - \sum_{k=1}^{3}\mathbb{A}_k\boldsymbol{v}_{x_k} + \mathbb{B}\boldsymbol{v} \equiv -L\boldsymbol{v} + 2\mathbb{B}\boldsymbol{v}$$

The linear operator L^* is clearly symmetric hyperbolic, like L.

Definition 4.6.3 *(weak solutions). A vector* $\boldsymbol{u}\in L^2(R_T)$ *is called a weak solution (with finite energy) of (4.82) and (4.91), with* $\boldsymbol{w}\in L^2(R_T)$), *if*

$$(4.93)\qquad ((L^*\boldsymbol{v},\boldsymbol{u})) = ((\boldsymbol{v},\boldsymbol{w}))$$

for every test vector $\boldsymbol{v}\in C^1_o(R_T)$ *and every* $T>0$.

This definition yields an actual generalization of the concept of solution. It does not require $\boldsymbol{u}$ and $\boldsymbol{w}$ to be smooth, it suffices that the scalar products which appear in eq. (4.93) be finite; in particular, $\boldsymbol{u}$ and $\boldsymbol{w}$ can be discontinuous. However, we have just seen that a smooth "classical" solution is also a weak solution, and conversely, every weak solution is also a classical solution if it is smooth enough (Exercise 10). Furthermore, a regularization theorem says that if $\boldsymbol{w}$ is smooth enough, then the weak solution is smooth too, and hence coincides with the classical solution (we omit the proof [8]).

Proposition 4.6.4 *Suppose* $\boldsymbol{v}(x_o,\boldsymbol{x}) \in C^1_o(R_T)$ *satisfies* $L^*\boldsymbol{v}=\mathbf{0}$ *in* R_T. *Then* $\boldsymbol{v}\equiv\mathbf{0}$ *in* R_T.

[8] see K.O. Friedrichs, Comm. Pure Appl. Math. 7 (1954), 345-392

Proof. By definition $\boldsymbol{v}(cT,\boldsymbol{x}) = \boldsymbol{0}$ and

$$L^*\boldsymbol{v} \equiv -L\boldsymbol{v} + 2\mathbb{B}\boldsymbol{v} \equiv -\boldsymbol{v}_{x_o} - \sum_{k=1}^{3} \mathbb{A}_k \boldsymbol{v}_{x_k} + \mathbb{B}\boldsymbol{v}$$

in R_T. Proceeding as in the proof of Proposition 4.6.2 we then find, for $0 < x_o < cT$

$$-\frac{1}{2}\frac{d}{dx_o}\int_{\mathbb{R}^3} |\boldsymbol{v}|^2\, d\boldsymbol{x} \;=-\int_{\mathbb{R}^3} \boldsymbol{v}\cdot\mathbb{B}\boldsymbol{v}\, d\boldsymbol{x} \;\leq\; 0$$

since the matrix $\mathbb{B}$ is positive semidefinite. Integrating over x_o from $c\tau$ to $cT,$ and applying the Schwartz inequality in $L^2(R_T)$ we obtain

$$\|\boldsymbol{v}(c\tau,\cdot)\|^2 \leq\; 2\,((\boldsymbol{v}, L^*\boldsymbol{v})) \;\leq\; 2\|\boldsymbol{v}\|_{T}\|L^*\boldsymbol{v}\|_{T} \qquad \forall\, \tau \in\; [0,T)$$

Suppose $\boldsymbol{v}$ is not identically zero in R_T. A further integration over τ from 0 to T then yields the inequality

$$0 < \|\boldsymbol{v}\|_{T} \leq 2cT\|L^*\boldsymbol{v}\|_{T} \tag{4.94}$$

which contradicts the assumption $L^*\boldsymbol{v}= 0$.

Proposition 4.6.5 *There is at least one* $\boldsymbol{u}\in L^2(R_T)$ *satisfying (4.93).*

Proof. If $\boldsymbol{v}$ and $\boldsymbol{v'}$ are in $C_o^1(R_T)$, we can define a new scalar product as

$$< \boldsymbol{v}, \boldsymbol{v'} >:= ((L^*\boldsymbol{v}, L^*\boldsymbol{v'}))$$

Indeed, this expression is bilinear symmetric, and the corresponding norm $|||\cdot|||$ (the "graph norm" of L^*)

$$|||\boldsymbol{v}|||^2 =< \boldsymbol{v}, \boldsymbol{v} >= ((L^*\boldsymbol{v}, L^*\boldsymbol{v})) = \|L^*\boldsymbol{v}\|_{T}^2$$

vanishes if and only if $\boldsymbol{v}$ is identically zero, since by force of (4.94)

$$\|\boldsymbol{v}\|_T^2 \leq (2cT)^2|||\boldsymbol{v}|||^2 \tag{4.95}$$

With this new scalar product $C_o^1(R_T)$ has the structure of a pre$-$Hilbert space. The completion of this pre-Hilbert space by means of Cauchy sequences with respect to the norm $|||\cdot|||$ yields a Hilbert space $\mathbb{H}$ [39], contained in $L^2(R_T)$. For all v$\in C_o^1(R_T)$ and w$\in L^2(R_T)$ the Schwartz inequality and eq. (4.95) yield

$$|((\boldsymbol{v}, \boldsymbol{w}))| \leq \|\boldsymbol{v}\|_T\|\boldsymbol{w}\|_T \leq K'_T\|\boldsymbol{w}\|_T|||\boldsymbol{v}||| \tag{4.96}$$

($K'_T > 0$). On the other hand, the definition of completion together with eq. (4.95) imply that for every $\boldsymbol{v} \in \mathbb{H}$ there is a Cauchy sequence $\boldsymbol{v}_n \in C^1_o(R_T)$ such that

$$\|\boldsymbol{v}_n - \boldsymbol{v}\|_T \leq 2cT|||\boldsymbol{v}_n - \boldsymbol{v}||| \to 0 \quad , \quad ((\boldsymbol{v}_n, \boldsymbol{w})) \to ((\boldsymbol{v}, \boldsymbol{w})) \quad , \quad |||\boldsymbol{v}_n||| \to |||\boldsymbol{v}|||$$

as $n \to +\infty$. The inequality (4.96) then holds for all $\boldsymbol{v} \in \mathbb{H}$ and $\boldsymbol{w} \in L^2(R_T)$ and implies that $((\mathbf{v}, \boldsymbol{w}))$ defines, for any $\boldsymbol{w} \in L^2(R_T)$, a continuous linear functional over $\mathbb{H}$. Riesz' representation theorem [2, 39] then says that the functional $((\boldsymbol{v}, \mathbf{w}))$ can be written as a scalar product

$$((\boldsymbol{v}, \boldsymbol{w})) = < \boldsymbol{v}, \boldsymbol{U} >$$

for some $\boldsymbol{U} = \boldsymbol{U}[\boldsymbol{w}] \in \mathbb{H}$. Since by definition $<\boldsymbol{v}, \boldsymbol{U}> = ((L^*\boldsymbol{v}, L^*\boldsymbol{U}))$ we obtain

$$((\boldsymbol{v}, \boldsymbol{w})) = ((L^*\boldsymbol{v}, L^*\boldsymbol{U}))$$

Eq. (4.93) becomes then $((L^*\boldsymbol{v}, \boldsymbol{u})) = ((L^*\boldsymbol{v}, L^*\boldsymbol{U}))$, that is,

$$((L^*\boldsymbol{v}, \boldsymbol{u} - L^*\boldsymbol{U})) = 0 \qquad \forall \boldsymbol{v} \in C^1_o(R_T)$$

with

$$\|L^*\boldsymbol{U}\|^2_T = ((L^*\boldsymbol{U}, L^*\boldsymbol{U})) = |||\boldsymbol{U}|||^2 < \infty$$

We conclude that $\boldsymbol{u} = L^*\boldsymbol{U}[\boldsymbol{w}]$ belongs to $L^2(R_T)$ and is a weak solution of (4.82) and (4.91).

Proposition 4.6.6 *The weak solution is unique.*

Proof. If $\boldsymbol{w} = 0$ eq. (4.93) shows that a weak solution $\boldsymbol{u} \in L^2(R_T)$ satisfies

$$((L^*\boldsymbol{v}, \boldsymbol{u})) = 0 \qquad \forall \boldsymbol{v} \in C^1_o(R_T) \tag{4.97}$$

All we need is to prove that $\boldsymbol{u} \equiv \mathbf{0}$ (a.e. in R_T). Consider the Cauchy problem

$$L^*\boldsymbol{v}_n = \boldsymbol{u}_n \in C^1_o(R_T) \quad , \qquad \boldsymbol{v}_n(0, \boldsymbol{x}) = 0$$

From Proposition 4.6.5 applied to L^* and the regularization theorem, the solution $\boldsymbol{v}_n \in C^1_o(R_T)$ exists for every $\boldsymbol{u}_n$. Let us choose $\boldsymbol{u}_n$ in such a way that $\|\boldsymbol{u}_n - \boldsymbol{u}\|_T \to 0$ as $n \to \infty$:this is possible because $C_o^1(R_T)$ is dense in $L^2(R_T)$ [13]. We have then

$$\|L^*\boldsymbol{v}_n - \boldsymbol{u}\|_T \to 0 \qquad as\ n \to \infty \tag{4.98}$$

and this equation implies that $\{\boldsymbol{v}_n\}$ is a Cauchy sequence in the norm $|||\cdot|||$:

$$\|L^*\boldsymbol{v}_n - L^*\boldsymbol{v}_m\|_T = |||\boldsymbol{v}_n - \boldsymbol{v}_m||| \to 0 \qquad as\ n, m \to \infty$$

As $\mathbb{H}$ is complete, there exists $\boldsymbol{v} \in \mathbb{H} \subset C_o^1(R_T)$ such that $|||\boldsymbol{v}_n - \boldsymbol{v}||| \to 0$ as $n \to \infty$, i.e.

$$\|L^*\boldsymbol{v}_n - L^*\boldsymbol{v}\|_T \to 0 \quad as\ n \to \infty \tag{4.99}$$

From eqs. (4.98), (4.99) it follows that $\boldsymbol{u} = L^*\boldsymbol{v}$, and from eq. (4.97) we have $\|\boldsymbol{u}\|_T = 0$. Therefore $\boldsymbol{u} \equiv \boldsymbol{0}$ (a.e. in R_T), and the weak solution is unique.

Remark 13. The initial class $L^2(\mathbb{R}^3)$ is persistent. It is possible to prove that the square summability of the derivatives of $\boldsymbol{u}$ is also a persistent property so that there is no loss of derivatives in the L^2 norm [2,18].

Remark 14. The proof of Proposition 4.6.5 is not constructive: a constructive existence proof can be obtained from the method of finite differences [31].

4.7 Characteristics and geometrical optics

4.7.1 Characteristics of the Maxwell equations.

Consider a weak solution $\boldsymbol{u}$ of the symmetric hyperbolic system

$$L\boldsymbol{u} := \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{u}_{x_\alpha} = \boldsymbol{w} \tag{4.100}$$

which for $\mathbb{A}_o = \mathbb{I}$, $\mathbb{B} = 0$ and constant matrices $\mathbb{A}_k$ defined by (4.83) and (4.84) coincides with the Maxwell system (4.82) in an isotropic homogeneous dielectric [9]. We have seen previously that the field $\boldsymbol{u}$, defined by

$$\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{X}) = (u_1, ..., u_6) := (\sqrt{\epsilon}\boldsymbol{E}(\boldsymbol{x}, t), \sqrt{\mu}\boldsymbol{H}(\boldsymbol{x}, t)) \tag{4.101}$$

[9] The extension to the case $\mathbb{B} \neq 0$ of conductors is straightforward

can be discontinuous. This is important, as the discontinuities of the electromagnetic field are the so-called signals: for example, a luminous signal separates a lighted region from a dark one and the separation surface carries in general some discontinuity of $\boldsymbol{u}$.

Suppose then that $\boldsymbol{u}$ is a piecewise smooth weak solution of (4.100). Precisely, let us assume that $\boldsymbol{u}$ is of class C^1 in R_T except for a discontinuity surface $\mathbb{S}$ through which $\boldsymbol{u}$ suffers a jump, so that the limits $\boldsymbol{u}_-$ and $\boldsymbol{u}_+$ on opposite $\pm$ sides of $\mathbb{S}$ exist and are finite. We suppose that the jump $[\boldsymbol{u}] = \boldsymbol{u}_+ - \boldsymbol{u}_-$ of $\boldsymbol{u}$ across $\mathbb{S}$ is a regular function of $\boldsymbol{X}$ on $\mathbb{S}$, and we denote by $\boldsymbol{\nu} = (\nu_o, ..., \nu_3) \equiv (\nu_o, \boldsymbol{N})$ the normal to $\mathbb{S}$ in space-time, oriented from the $-$side to the $+$side. In §1.4 we have considered the case of a discontinuity surface fixed in space, i.e. independent of time. In this section we examine the general case when $\mathbb{S}$ is a four-dimensional surface in space-time with cartesian coordinates $\boldsymbol{X} = (x_o, ..., x_3)$, which can be viewed as a time-dependent surface in ordinary space.

The definition (4.93) of weak solution, with $L^* = -L$, yields $-((L\,\boldsymbol{v}, \boldsymbol{u})) = ((\boldsymbol{v}, \boldsymbol{w}))$, i.e.

$$- \int_{R_T} \boldsymbol{u} \cdot \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{v}_{x_\alpha} d\boldsymbol{X} \; = \int_{R_T} \boldsymbol{v} \cdot \boldsymbol{w} \, d\boldsymbol{X} \tag{4.102}$$

for every test vector $\boldsymbol{v} \in C^1_o(R_T)$. Since by assumption $\boldsymbol{u}$ is of class C^1 in the two regions D_- and D_+ in which $\mathbb{S}$ divides R_T , taking a test vector $\boldsymbol{v}$ in eq. (4.102) with support in D_- or in D_+ , and by applying the Gauss Lemma backwards we see that $\boldsymbol{u}$ satisfies the equation

$$\int_D \boldsymbol{v} \cdot (\boldsymbol{w} - \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{u}_{x_\alpha}) d\boldsymbol{X} = 0 \tag{4.103}$$

where D is the open set $D_+ \cup D_-$. This equation implies that $L\,\boldsymbol{u} = \boldsymbol{w}$ in D. If on the other hand $\boldsymbol{v}$ is a generic test vector with support in R_T, since $\boldsymbol{u}$ is piecewise smooth the Gauss Lemma can be applied separately in D_- and D_+ and from eq. (4.102) we find

$$0 = \int_{R_T} (\boldsymbol{u} \cdot \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{v}_{x_\alpha} + \boldsymbol{v} \cdot \boldsymbol{w}) \, d\boldsymbol{X} \; \equiv \int_D (\boldsymbol{u} \cdot \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{v}_{x_\alpha} + \boldsymbol{v} \cdot \boldsymbol{w}) \, d\boldsymbol{X}$$

$$= \int_D \boldsymbol{v} \cdot (\boldsymbol{w} - \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{u}_{x_\alpha}) d\boldsymbol{X} + \int_{\mathbb{S}} \boldsymbol{v} \cdot \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{\nu}_\alpha [\boldsymbol{u}] \, d\mathbb{S}$$

and substituting into eq. (4.103) we finally obtain

$$\int_{\mathbb{S}} \boldsymbol{v} \cdot \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \boldsymbol{\nu}_\alpha [\boldsymbol{u}] \, d\mathbb{S} = 0$$

As the test vector $\boldsymbol{v}$ is arbitrary, the linear algebraic system

$$\mathbb{A}(\boldsymbol{\nu}(\boldsymbol{X}))[\boldsymbol{u}] = \mathbf{0} \tag{4.104}$$

must be satisfied for all $\boldsymbol{X} \in \mathbb{S}$, where $\mathbb{A}(\boldsymbol{X})$ is the characteristic matrix , defined by

$$\mathbb{A} = \mathbb{A}(\boldsymbol{\nu}(\boldsymbol{X})) := \sum_{\alpha=0}^{3} \mathbb{A}_\alpha \nu_\alpha(\boldsymbol{X}) \tag{4.105}$$

and $\boldsymbol{\nu}(\boldsymbol{X})$ is the 4D normal at the point $\boldsymbol{X} \in \mathbb{S}$. The homogeneous algebraic system (4.104) puts a restriction on $\boldsymbol{\nu}(\boldsymbol{X})$: indeed, if $[\boldsymbol{u}] \neq \mathbf{0}$, as is the case here, the characteristic equation

$$det\, \mathbb{A}(\boldsymbol{\nu}(\boldsymbol{X})) = 0 \tag{4.106}$$

must be satisfied at all points of $\mathbb{S}$.

Definition 4.7.1 *A surface $\mathbb{S}$ in spacetime $\mathbb{R}^4$ is called a characteristic (or characteristic surface) of the hyperbolic system (4.100) if it satisfies (4.106). Viewed in ordinary space $\mathbb{R}^3$, a characteristic is called a wavefront for every fixed $x_o = ct$.*

We conclude that a piecewise smooth vector function $\boldsymbol{u}$ is a weak solution of (4.100) only if the surface $\mathbb{S}$ across which the jump discontinuity $[\boldsymbol{u}] \neq \mathbf{0}$ is a characteristic. In other words, the jump discontinuities of $\boldsymbol{u}$ are localized on characteristics in space-time, and the signals propagate as wavefronts in ordinary space. The existence of wavefronts is guaranteed by the hyperbolicity of the system (4.100) [18].

What are the wavefronts for the Maxwell equations and their possible propagation speeds? Suppose that the characteristic surface $\mathbb{S}$ is given by the implicit equation

$$\Phi(x_o, \boldsymbol{x}) = 0 \tag{4.107}$$

with Φ a smooth function satisfying [10]

$$\Phi^2_{x_o} + |grad_{\boldsymbol{x}} \Phi|^2 > 0 \tag{4.108}$$

in a neighborhood of $\mathbb{S}$. The four-dimensional normal $\boldsymbol{\nu} = (\nu_o, \boldsymbol{N})$ to $\mathbb{S}$ is then defined by

$$\nu_o = \frac{\Phi_{x_o}}{\sqrt{\Phi^2_{x_o} + |grad_{\boldsymbol{x}} \Phi|^2}}, \quad \boldsymbol{N} = \frac{grad_{\boldsymbol{x}} \Phi}{\sqrt{\Phi^2_{x_o} + |grad_{\boldsymbol{x}} \Phi|^2}} \tag{4.109}$$

The characteristic equation (for $\mathbb{A}_o = \mathbb{I}$)

$$det(\nu_o \mathbb{I} + \sum_{k=1}^{3} \mathbb{A}_k \nu_k(\boldsymbol{X})) = 0 \tag{4.110}$$

can be viewed as an algebraic equation of 6th degree in the unknown ν_o for a given 3D vector $\boldsymbol{N} := (\nu_1, \nu_2, \nu_3)$. Since the 6×6 real matrix

$$\sum_{k=1}^{3} \mathbb{A}_k \nu_k(\boldsymbol{X})$$

is symmetric for any $\boldsymbol{X} \in \mathbb{S}$, equation (4.110) has 6 real roots for $\nu_o = \nu_o(\boldsymbol{N})$, not necessarily distinct. Precisely, from the definitions (4.83) and (4.84) the characteristic matrix for the Maxwell system turns out to be

$$\mathbb{A} = \nu_o \mathbb{I} + \sum_{k=1}^{3} \mathbb{A}_k \nu_k(\boldsymbol{X}) = \begin{pmatrix} \nu_o & 0 & 0 & 0 & \nu_3 & -\nu_2 \\ 0 & \nu_o & 0 & -\nu_3 & 0 & \nu_1 \\ 0 & 0 & \nu_o & \nu_2 & -\nu_1 & 0 \\ 0 & -\nu_3 & \nu_2 & \nu_o & 0 & 0 \\ \nu_3 & 0 & -\nu_1 & 0 & \nu_o & 0 \\ -\nu_2 & \nu_1 & 0 & 0 & 0 & \nu_o \end{pmatrix}$$

and the characteristic equation for the Maxwell system is given by

$$det\Big[\mathbb{A}(\boldsymbol{\nu}(\boldsymbol{X}))\Big] \equiv \nu_o^2(\nu_o^2 - |\boldsymbol{N}|^2)^2 = 0 \tag{4.111}$$

where $\boldsymbol{N}|^2 = \nu_1^2 + \nu_2^2 + \nu_3^2$ (Exercise 11). Hence there are three double roots

$$\nu_o = 0 \,, \quad \nu_o = |\boldsymbol{N}| \,, \quad \nu_o = -|\boldsymbol{N}|$$

[10] this condition is required by the implicit function theorem [36]

depending only on $|\boldsymbol{N}|$. Eq. (4.109) yields

$$|\boldsymbol{N}|\Phi_{x_o} = \nu_o|grad_{\boldsymbol{x}}\Phi| \tag{4.112}$$

and for $\nu_o = 0, |\boldsymbol{N}|, -|\boldsymbol{N}|$ we obtain three corresponding partial differential equations for the function $\Phi(\boldsymbol{X}), \boldsymbol{X} = (x_o, \boldsymbol{x}) \in \mathbb{R}^4$, whose solutions substituted into (4.107) define three families of wavefronts for the Maxwell equations. In order to find the physical interpretation of this, it is convenient to revert to the time variable $t = x_o/c$ and to remark that the 3D normal $\boldsymbol{n} = \boldsymbol{n}(\boldsymbol{X})$ to the wavefront

$$\boldsymbol{n}(\boldsymbol{X}) = \frac{grad_{\boldsymbol{x}}\Phi}{|grad_{\boldsymbol{x}}\Phi|} \tag{4.113}$$

is well defined, since eqs. (4.108) implies that

$$|grad_{\boldsymbol{x}}\Phi| > 0$$

in a neighborhood of $\mathbb{S}$, for any fixed $x_o = ct$. Eq. (4.109) implies that

$$\nu_o = \frac{|\boldsymbol{N}|\ \Phi_{x_o}}{|grad_{\boldsymbol{x}}\Phi|}, \quad \boldsymbol{N} = |\boldsymbol{N}|\,\boldsymbol{n}$$

Differentiating eq.(4.107), with $x_o = ct$, we obtain

$$\Phi_{x_o} dx_o + grad_{\boldsymbol{x}}\Phi \cdot d\boldsymbol{x} \equiv c\Phi_{x_o} dt + |grad_{\boldsymbol{x}}\Phi|\,\boldsymbol{n} \cdot d\boldsymbol{x} = 0$$

and this equation shows that the quantity

$$s(\boldsymbol{X}) := \boldsymbol{n} \cdot \frac{d\boldsymbol{x}}{dt} = -\frac{c\Phi_{x_o}}{|grad_{\boldsymbol{x}}\Phi|} \equiv -\frac{\Phi_t}{|grad_{\boldsymbol{x}}\Phi|}$$

is the normal speed of propagation, or characteristic speed, of the wavefront. As $x_o = ct$ increases, every point $\boldsymbol{x}$ of the wavefront moves with velocity $s(\boldsymbol{X})\boldsymbol{n}(\boldsymbol{X})$ in $\mathbb{R}^3$.

By force of the preceding relations $-s$ is proportional to Φ_{x_o} (hence to ν_o) and eqs. (4.111), (4.112) can be written as

$$s^2(s^2 - c^2)^2 = 0 \tag{4.114}$$

$$\Phi_t(t, \boldsymbol{x}) = s|grad_{\boldsymbol{x}}\Phi(t, \boldsymbol{x})| \ , \quad (t, \boldsymbol{x}) \in \mathbb{R}^4 \tag{4.115}$$

By taking s equal to the appropriate root

$$s = 0 \quad ; \quad s = c \quad ; \quad s = -c \tag{4.116}$$

of (4.114) we obtain three families of wavefronts

$$\Phi_t = 0 \ , \quad \Phi_t(t, \boldsymbol{x}) = c|grad_{\boldsymbol{x}}\Phi(t, \boldsymbol{x})| \ , \quad \Phi_t(t, \boldsymbol{x}) = -c|grad_{\boldsymbol{x}}\Phi(t, \boldsymbol{x})|$$

across which the jump discontinuity $[\boldsymbol{u}]$ satisfies

$$s\epsilon[\boldsymbol{E}] + \boldsymbol{n} \wedge [\boldsymbol{H}] = \boldsymbol{0} \quad , \quad s\mu[\boldsymbol{H}] - \boldsymbol{n} \wedge [\boldsymbol{E}] = \boldsymbol{0} \tag{4.117}$$

with $s = 0$, $\pm c$ (Exercise 12).

Summarizing, for every choice of the unit vector $\boldsymbol{n}$, the normal to the wavefront in ordinary space, there are six real characteristic speeds s, counted with their multiplicity. A system (4.100) with this property, which is a direct consequence of the symmetry of the matrices $\mathbb{A}_\alpha$, is called hyperbolic. The Maxwell system, however, is not strictly hyperbolic, since the six roots coincide in pairs and yield only three distinct characteristic speeds, that is three families of wavefronts with three distinct normal propagation speeds $0, +c, -c$. Let us examine the three cases in more detail.

(i) : $\nu_o = s = 0$ (contact discontinuity). Φ satisfies

$$\Phi_t(t, \boldsymbol{x}) = 0 \ , \quad (t, \boldsymbol{x}) \in \mathbb{R}^4$$

the characteristics are arbitrary time-independent surfaces of equation $\Phi(\boldsymbol{x}) = 0, \boldsymbol{x} \in \mathbb{R}^3$ and across these fixed surfaces the jumps of $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy

$$\boldsymbol{n} \wedge [\boldsymbol{H}] = \boldsymbol{n} \wedge [\boldsymbol{E}] = \boldsymbol{0}$$

These are the jump surfaces $\mathbb{S}$ already considered in §1.4.

(ii), (iii) : $\nu_o = \pm|\boldsymbol{N}|$, $s = \pm c$ (linear shocks). Φ satisfies

$$\Phi_t = \pm c|grad\,\Phi(t, \boldsymbol{x})| \ , \quad (t, \boldsymbol{x}) \in \mathbb{R}^4 \tag{4.118}$$

and the wavefronts propagate in ordinary space with the normal speed $\pm c$. These characteristics are the same as for the wave equation [18]. In particular, the surfaces of constant phase of the plane waves (4.10) are wavefronts. Eqs. (4.117) become

$$\pm c\epsilon[\boldsymbol{E}] + \boldsymbol{n} \wedge [\boldsymbol{H}] = \boldsymbol{0}, \qquad \pm\, c\mu[\boldsymbol{H}] - \boldsymbol{n} \wedge [\boldsymbol{E}] = \boldsymbol{0}$$

so that the vectors $[\boldsymbol{E}], [\boldsymbol{H}], \boldsymbol{n}$ are mutually orthogonal

$$[\boldsymbol{E}] \cdot \boldsymbol{n} = [\boldsymbol{H}] \cdot \boldsymbol{n} = [\boldsymbol{E}] \cdot [\boldsymbol{H}] = 0$$

It follows that the normal components of $\boldsymbol{E}$ and $\boldsymbol{H}$ are continuous across $\mathbb{S}$, the jumps of $\boldsymbol{E}$ and $\boldsymbol{H}$ are tangential to $\mathbb{S}$, $[\boldsymbol{E}] \wedge [\boldsymbol{H}]$ is parallel to $\boldsymbol{n}$, and $\sqrt{\epsilon}\,|[\boldsymbol{E}]| = \sqrt{\mu}\,|[\boldsymbol{H}]|$, so that $[\boldsymbol{E}], [\boldsymbol{H}]$ can only vanish simultaneously. Since as already remarked $\Phi_t \neq 0$ in a neighborhood of $\mathbb{S}$, the characteristics can be written locally in the explicit form

$$\mathcal{I}(\boldsymbol{x}) = \pm\, ct + \text{constant}$$

and we can take a generating function Φ linear in x_o

$$\Phi(x_o, \boldsymbol{x}) = \mathcal{I}(\boldsymbol{x}) \mp x_o + \text{constant} \tag{4.119}$$

The function $\mathcal{I}(\boldsymbol{x})$, called the "eikonal" function , satisfies then the eikonal equation [11]

$$|grad\,\mathcal{I}| = 1 \tag{4.120}$$

and the normal to the wavefronts is given by $\boldsymbol{n} = grad\,\mathcal{I}$.

In one space variable x eq. (4.118) becomes $\Phi_t = \pm c\,\Phi_x$ that is, $dx/dt = \pm c$. Thus the one-dimensional Maxwell equations have characteristic curves given by the two families of straight lines

$$x \pm ct = constant$$

as for the vibrating string equation [2].

4.7.2 High frequency approximation.

The concept of characteristic and the propagation of discontinuities is strictly related to a high frequency approximation known as geometrical optics. We have seen that the determinant of the characteristic matrix is given by $\nu_o^2 Q^2(\boldsymbol{\nu})$, where, from eq. (4.111),

$$Q(\boldsymbol{\nu}) := \nu_o^2 - \sum_{k=1}^{3} \nu_k^2 \tag{4.121}$$

[11] if we put $\Phi = c_o t - \mathcal{I}(\boldsymbol{x})$+constant the eikonal equation becomes $|grad\,\mathcal{I}|^2 = n_r$, where $n_r = c_o/c$ is the refractive index of the medium

Every characteristic surface $\Phi(x_o,\boldsymbol{x}) = 0$ corresponding to the two characteristic speeds $s = \pm c$, i.e. satisfying $Q(\boldsymbol{\nu}) = 0$, is generated by bicharacteristic rays $\boldsymbol{X}=\boldsymbol{X}(\tau)$, $\boldsymbol{\nu} = \boldsymbol{\nu}(\tau)$ defined in terms of a curvilinear coordinate τ by the canonic system [12]

$$\frac{dX_\alpha}{d\tau} = \frac{\partial Q}{\partial \nu_\alpha} \;,\; \frac{d\boldsymbol{\nu}_\alpha}{d\tau} = -\frac{\partial Q}{\partial X_\alpha} \qquad (\alpha = 0, ..., 3) \tag{4.122}$$

with $Q = 0$. Since Q is independent of $\boldsymbol{X}$

$$\boldsymbol{\nu} \cdot \frac{d\boldsymbol{X}}{d\tau} = 2Q = 0$$

it follows that $\boldsymbol{\nu}$ is constant along each ray

$$\frac{d\nu_\alpha}{d\tau} = 0 \qquad \Rightarrow \quad \nu_\alpha(\boldsymbol{X}(\tau)) = \text{constant} \qquad (\alpha = 0, ..., 3)$$

and the rays lie on characteristic surfaces $\Phi(x_o,\boldsymbol{x}) = 0$ in $\mathbb{R}^4$. In fact, the characteristic surfaces are generated by bicharacteristic rays [18]. Eqs. (4.121) and (4.122) imply that

$$\frac{dx_o}{d\tau} = 2\nu_o \;,\; \frac{dx_k}{d\tau} = -2\nu_k \qquad (k = 1, 2, 3)$$

and by eliminating τ and taking eq. (4.113) and ff. into account we obtain

$$\frac{dx_k}{dx_o} = -\frac{\nu_k}{\nu_o} = \pm n_k \qquad (k = 1, 2, 3)$$

where $x_o = ct$ and $\boldsymbol{n}= grad\,\mathcal{I}$ is the normal to the wavefronts in ordinary space. Thus the bicharacteristic rays for the Maxwell equations are given by the family of straight lines in space-time

$$\boldsymbol{x} - \boldsymbol{x}_o = \pm c\boldsymbol{n}_o t$$

where $\boldsymbol{n}_o$ is the (arbitrary) value of $grad\,\mathcal{I}$ at the initial point of the ray. Interpreted in the ordinary $\mathbb{R}^3$ space, the rays are arbitrary straight trajectories, orthogonal to the wavefronts, on which the point $\boldsymbol{x}$ moves with speed c (Exercise 13).

[12] in the theory of first order PDE's, the canonic system (4.122) defines the "bicharacteristic strips" for the nonlinear Hamilton-Jacobi equation $Q(\Phi_{x_o}, grad\,\Phi) = 0$ [18]

The bicharacteristic rays thus defined correspond to the intuitive concept of luminous rays in (geometrical) optics. In order to see this, let us seek solutions of the Maxwell equations

$$\epsilon\frac{\partial \boldsymbol{E}}{\partial t} = curl\ \ \boldsymbol{H}\ ,\ \ \mu\frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E}$$

(that is of the system (4.100) with $\boldsymbol{w}=\boldsymbol{0}$) of the form

$$\boldsymbol{E} = \boldsymbol{e}(\boldsymbol{x})e^{i\omega(t-\,\mathcal{I}(\boldsymbol{x})/c)} \quad , \qquad \boldsymbol{H} = \boldsymbol{h}(\boldsymbol{x})e^{i\omega(t-\,\mathcal{I}(\boldsymbol{x})/c)} \tag{4.123}$$

where the amplitude vectors $\boldsymbol{e}$ and $\boldsymbol{h}$ are bounded functions of $\boldsymbol{x}$ together with their first partial derivatives. Letting $\omega \to \infty$ we obtain

$$\begin{aligned} i\omega\mu\boldsymbol{h} &= -curl\,\boldsymbol{e} + i\omega c^{-1} grad\,\mathcal{I}\wedge \boldsymbol{e} \ \cong\ i\omega c^{-1} grad\,\mathcal{I}\wedge \boldsymbol{e}\\ i\omega\epsilon\boldsymbol{e} &= curl\,\boldsymbol{h} - i\omega c^{-1} grad\,\mathcal{I}\wedge \boldsymbol{h} \cong\ -\ i\omega c^{-1} grad\,\mathcal{I}\wedge \boldsymbol{h} \end{aligned}$$

(cfr. §4.2). We will see shortly that the function $\mathcal{I}$ is indeed the eikonal, and therefore defines the wavefronts $\mathcal{I}(\boldsymbol{x}) = \pm ct$+constant. The amplitudes $\mathbf{e}$ and $\boldsymbol{h}$ then satisfy as $\omega \to \infty$ the same relations

$$c\mu\boldsymbol{h}\ \ \cong\ \ \pm\ \ \boldsymbol{n}\wedge\boldsymbol{e}\quad ,\qquad c\epsilon\mathbf{e}\cong\ \ \mp\,\boldsymbol{n}\wedge\boldsymbol{h}$$

found in eq. (4.117) for the jumps $[\boldsymbol{E}]$ and $[\boldsymbol{H}]$, with $s = \pm c$. It follows that $\boldsymbol{e}$, $\boldsymbol{n}$, $\boldsymbol{h}$ are mutually orthogonal, the electromagnetic wave is transversal and the amplitudes $\boldsymbol{e}(\boldsymbol{x})$ and $\boldsymbol{h}(\boldsymbol{x})$ behave like discontinuities of the field across wavefronts. Eliminating $\boldsymbol{h}$ yields

$$\boldsymbol{e}\ \ \cong\ \ -\frac{1}{\epsilon\mu c^2} grad\,\mathcal{I}\wedge(grad\,\mathcal{I}\wedge\boldsymbol{e}) = |grad\,\mathcal{I}|^2\boldsymbol{e}$$

whence $|grad\,\mathcal{I}| = 1$. Therefore $\mathcal{I}$ satisfies the eikonal equation (4.120), and the surfaces of constant phase for $\boldsymbol{E}$ and $\boldsymbol{H}$ are wavefronts (cf. (4.123)).

We conclude that, in the approximation of geometrical optics, the electromagnetic wave is replaced by the rays, which propagate in the direction $\boldsymbol{n}$ orthogonal to the wavefronts. The ray direction $\boldsymbol{n}$ coincides with the direction of the Poynting vector $\boldsymbol{E}\wedge\boldsymbol{H}$, and the vector

$$\frac{\boldsymbol{E}\wedge\boldsymbol{H}}{\boldsymbol{E}\ \cdot\boldsymbol{D}}$$

yields the energy velocity along the rays. On the wavefronts $\boldsymbol{E}$ and $\boldsymbol{H}$ have constant phases and slowly varying amplitudes $\boldsymbol{e}$ and $\boldsymbol{h}$, which behave like jump discontinuities of the field across the fronts. The wave propagation phenomena in this approximation are described in terms of ray geometry, and the Maxwell equations are replaced by ordinary differential equations which govern the ray geometry and the evolution of discontinuities along the rays. The approximation of geometrical optics fails where the amplitudes or the eikonal vary brusquely, like in the presence of diffraction or focussing phenomena.

4.8 Reflection, refraction and Snell's law. Total reflection.

4.8.1 Snell's law.

Suppose that two homogeneous dielectrics, with different permittivities $\epsilon_1\,(x_3 < 0)$, $\epsilon_2\,(x_3 > 0)$ and the same permeability μ_o, are separated by the (x_1,x_2)-plane $\mathbb{S}$, and that a given linearly polarized plane monochromatic wave

$$\boldsymbol{E}_i = \boldsymbol{E}_o e^{i\omega\,(t-\boldsymbol{k}\cdot\boldsymbol{x}/c_1)}\ ,\quad \boldsymbol{H}_i = \boldsymbol{H}_o e^{i\omega(t-\boldsymbol{k}\cdot\boldsymbol{x}/c_1)} \qquad (\boldsymbol{x} = (x_1, x_2, x_3))$$

is obliquely incident on $\mathbb{S}$ in the first medium, $x_3 < 0$. According to eqs. (4.10) and (4.16) this incoming wave has wavenumber $\boldsymbol{p}= \omega\boldsymbol{k}/c_1$, where $c_1 = (\epsilon_1\mu_o)^{-1/2}$ and

$$\boldsymbol{k} = \mathbf{c}_3 cos\,\theta_i + \mathbf{c}_2 sin\,\theta_i$$

is contained in the (x_2, x_3)-plane (the incidence plane) and forms an incidence angle θ_i with the positive x_3- axis. The real amplitude vectors $\boldsymbol{E}_o, \boldsymbol{H}_o$ satisfy the orthogonality relations

$$\boldsymbol{k}\cdot\boldsymbol{E}_o = \boldsymbol{k}\cdot\boldsymbol{H}_o = \boldsymbol{E}_o\cdot\boldsymbol{H}_o = 0$$

and by virtue of eq. (4.14) $\boldsymbol{H}_i$ is known if $\boldsymbol{E}_i is$:

$$\boldsymbol{H}_i = \frac{1}{c_1\mu_o}\ \boldsymbol{k}\wedge\boldsymbol{E}_i \tag{4.124}$$

Let $0 \leq \theta_i < \pi/2$ and $\boldsymbol{n}=\mathbf{c}_3$ denote the normal to $\mathbb{S}$. Then $\boldsymbol{n}\wedge\boldsymbol{E}_i, \boldsymbol{n}\wedge\boldsymbol{H}_i$ do not vanish on $\mathbb{S}$ and, since they must be continuous the second dielectric

$x_3 > 0$ must contain a transmitted wave which can be sought in the shape of a plane monochromatic wave

$$\boldsymbol{E}_t = \boldsymbol{E}'_o\, e^{i\omega'(t-\boldsymbol{k}'\cdot\boldsymbol{x}/c_2)} \quad , \quad \boldsymbol{H}_t = \boldsymbol{H}'_o e^{i\omega'(t-\boldsymbol{k}'\cdot\boldsymbol{x}/c_2)}$$

with wavenumber $\boldsymbol{p}= \omega\boldsymbol{k}'/c_2$, where $c_2 = (\epsilon_2\mu_o)^{-1/2}$ and

$$\boldsymbol{k}' = \mathbf{c}_3 cos\,\theta_t + \mathbf{c}_2 sin\,\theta_t$$

is also contained in the incidence plane (x_2, x_3) and forms an angle θ_t with the positive x_3-axis. The vectors $\boldsymbol{E}'$, $\boldsymbol{H}', \boldsymbol{k}'$ are mutually orthogonal, and the relation between $\boldsymbol{H}_t$ and $\boldsymbol{E}_t$ is now

(4.125) $$\boldsymbol{H}_t = \frac{1}{c_2\mu_o}\ \boldsymbol{k}' \wedge \boldsymbol{E}_t$$

The transmitted wave is also called *refracted* wave, especially at oblique incidence, and θ_t is the refraction angle. At the interface $x_3 = 0$, $\boldsymbol{n}\wedge\boldsymbol{E}_i$ must match continuously with $\boldsymbol{n}\wedge\boldsymbol{E}_t$

$$\boldsymbol{n} \wedge \boldsymbol{E}_o\ exp\Big[i\omega(t - \frac{x_2 sin\,\theta_i}{c_1})\Big] = \ \boldsymbol{n} \wedge \boldsymbol{E}'_o\ exp\Big[i\omega'(t - \frac{x_2 sin\,\theta_t}{c_2})\Big]$$

for all t, x_2. This is clearly impossible unless

$$\omega' = \omega$$

and

$$\frac{sin\,\theta_i}{c_1} = \frac{sin\,\theta_t}{c_2}$$

The first equation says that the frequency does not change [13], the second is Snell's refraction law. If these two conditions are satisfied, the phases of the incident and transmitted waves coincide at the interface $x_3 = 0$. It remains to match the amplitudes of the electric and magnetic fields: as we will see in a particular case, eqs. (4.124) and (4.125) imply that this is possible only if a second wave, called reflected wave

$$\boldsymbol{E}_r = \boldsymbol{E}''_o\, e^{i\omega\,(t-\boldsymbol{k}''\cdot\boldsymbol{x}/c_1)} \quad , \quad \boldsymbol{H}_r = \boldsymbol{H}''_o\, e^{i\omega\,(t-\boldsymbol{k}''\cdot\boldsymbol{x}/c_1)}$$

[13] hence the wavelengths $2\pi c_1/\omega$, $2\pi c_2/\omega$ are different in the two media

is present in the first medium ($x_3 < 0$), with

$$\boldsymbol{k}'' = -\mathbf{c}_3 cos\,\theta_r + \mathbf{c}_2 sin\,\theta_r$$

where

$$\theta_r = \pi - \theta_i \tag{4.126}$$

and

$$\boldsymbol{H}_r = \frac{1}{c_1\mu_o}\ \boldsymbol{k}\wedge\boldsymbol{E}_r \tag{4.127}$$

In this way $sin\theta_r = sin\theta_i$ and the phases of the three waves coincide for $x_3 = 0$. Moreover the matching of the amplitudes can be carried out only if the vectors $\boldsymbol{k}$,$\boldsymbol{k}'$ and $\boldsymbol{k}''$ are coplanar, so that the planes of reflection ($\boldsymbol{k}''$,$\mathbf{c}_3$) and transmission ($\boldsymbol{k}'$,$\mathbf{c}_3$) coincide with the plane of incidence ($\boldsymbol{k}$,$\mathbf{c}_3$).

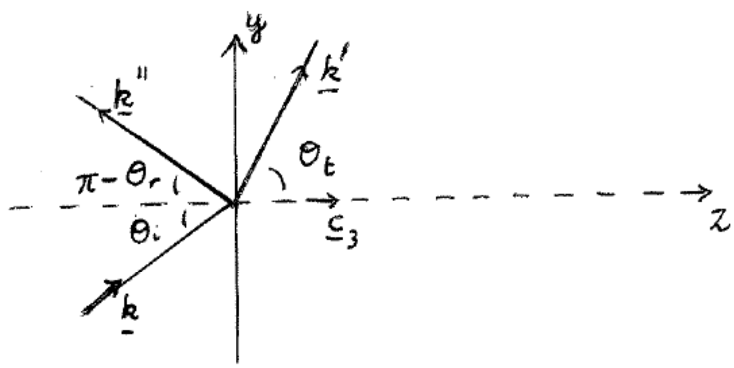

Figure 4.2: Geometry in Snell's law

Equation (4.126) says that the angles of incidence and reflection coincide if referred to the negative x_3-axis, and Snell's refraction law can be written as

$$n_i sin\,\theta_i = n_t\,sin\,\theta_t \tag{4.128}$$

where $n_i = n_1 = \sqrt{\epsilon_1/\epsilon_o}$, $n_t = n_2 = \sqrt{\epsilon_2/\epsilon_o}$ are the refractive indices of the two dielectrics.

4.8.2 Total reflection and evanescent waves.

At grazing incidence the refracted wave in a medium of lower permittivity becomes an evanescent wave. Suppose $\epsilon_1 > \epsilon_2$, that is $n_i > n_t$, and let

$$\theta_i \;=\; arcsin\frac{n_t}{n_i}$$

Then the angle θ_t defined by Snell's law (4.128)

$$\theta_t = arcsin(\frac{n_i}{n_t} sin\,\theta_i)$$

is equal to $\frac{\pi}{2}$, and the refracted wave propagates along the interface. If

$$arcsin\frac{n_t}{n_i} \; < \theta_i < \frac{\pi}{2} \tag{4.129}$$

there exists no real angle θ_t defined by Snell's law. In this case Snell's law can be satisfied by taking a complex angle θ_t with real part $\frac{\pi}{2}$ and imaginary part β

$$\theta_t = \frac{\pi}{2} + i\beta$$

where $\beta \in \mathbb{R}$ is a function of θ_i defined by

$$\beta := cosh^{-1}(\frac{c_2}{c_1} sin\,\theta_i) \equiv cosh^{-1}(\frac{n_i}{n_t} sin\,\theta_i)$$

In this way the complex sine of θ_t is real and greater than one, while the complex cosine is imaginary:

$$\begin{aligned} sin\,\theta_t &= sin(\frac{\pi}{2} + i\beta) = cos\,(i\beta) = cosh\,\beta \\ cos\,\theta_t &= cos\,(\frac{\pi}{2} + i\beta) = -sin\,(i\beta) = -isinh\,\beta \end{aligned}$$

and Snell's law is satisfied:

$$n_i sin\,\theta_i - n_t\, sin\,\theta_t = n_i sin\,\theta_i - n_t\, cosh\,\beta = n_i sin\,\theta_i - n_i\, sin\,\theta_i = 0$$

Since

$$sin\,\theta_t = \frac{n_i}{n_t} sin\,\theta_i \quad , \quad cos\,\theta_t = \; -i\left[\frac{n_i^2}{n_t^2} sin^2\theta_i - 1\right]^{\frac{1}{2}}$$

the corresponding wavenumber $\boldsymbol{k}' = \mathbf{c}_3 cos\,\theta_t + \mathbf{c}_2 sin\,\theta_t$ is also complex:

$$\boldsymbol{k}' = \mathbf{c}_2\, cosh\,\beta - i\mathbf{c}_3\, sinh\,\beta = \mathbf{c}_2 \frac{n_i}{n_t}\, sin\,\theta_i - i\mathbf{c}_3 \left[\frac{n_i^2}{n_t^2} sin^2\theta_i - 1 \right]^{\frac{1}{2}}$$

with a real part parallel to $\mathbf{c}_2$ and an imaginary part parallel to $\mathbf{c}_3$. If we put, in the notations of §4.2.1,

$$\text{(4.130)} \quad \boldsymbol{P}' := \frac{1}{c_1} \omega sin\,\theta_i\, \mathbf{c}_2 \equiv P'\mathbf{c}_2\ , \quad \boldsymbol{p}' := \frac{\omega}{c_2} \left[\frac{n_i^2}{n_t^2} sin^2\theta_i - 1 \right]^{\frac{1}{2}} \mathbf{c}_3\ \equiv p'\mathbf{c}_3$$

we have

$$\frac{\omega}{c_2}\boldsymbol{k}' = \boldsymbol{P}' - i\boldsymbol{p}' \equiv P'\mathbf{c}_2 - ip'\mathbf{c}_3$$

with $\boldsymbol{p}' \cdot \boldsymbol{P}' = 0$.

Summarizing, the electric field of the refracted wave ($x_3 > 0$) under condition (3.129) is

$$\text{(4.131)} \qquad \boldsymbol{E}_t = \boldsymbol{E}_o' e^{-\,\boldsymbol{p}'\cdot\boldsymbol{x}}\, e^{i(\omega t - \boldsymbol{P}'\cdot\boldsymbol{x})} \ \equiv\ \boldsymbol{E}_o'\, e^{-p'x_3}\, e^{i(\omega t - P'x_2)}$$

This is an evanescent wave, propagating along the interface in the direction of the x_2- axis, and damped orthogonally to the interface in the positive direction of the x_3-axis. Eq. (4.130) shows that $\boldsymbol{p}'$ is completely determined by the incidence angle via Snell's law and is proportional to ω, so that also

$$|\boldsymbol{P}'| = \sqrt{\ |\boldsymbol{p}'|^2 + \frac{\omega^2}{c_2^2}\ } = \sqrt{\frac{\omega^2}{c_1^2} sin^2\,\theta_i} = \frac{\omega}{c_1} |sin\ \theta_i|$$

is proportional to ω, and satisfies

$$|\boldsymbol{P}'|^2 - |\boldsymbol{p}'|^2 = \frac{\omega^2}{c_2^2}$$

This evanescent wave is therefore non-dispersive (see §4.2.1). It is also non-transversal. Indeed, the magnetic field of the refracted wave, given by eqs. (4.125), (4.130) and (4.131), is of the form

$$\boldsymbol{H}_t = \frac{1}{c_2\mu_o}\ \boldsymbol{k}' \wedge \boldsymbol{E}_o'\, e^{-p'x_3}\, e^{i(\omega t - P'x_2)}$$

where

$$\boldsymbol{k}' \wedge \boldsymbol{E}'_o = \ \mathbf{c}_2 \wedge \boldsymbol{E}'_o \, cosh\,\beta \ - i \ \mathbf{c}_3 \wedge \boldsymbol{E}'_o sinh\,\beta \ , \ \ \boldsymbol{k}' \cdot \boldsymbol{E}'_o \ = 0$$

Therefore if $\boldsymbol{E}_t$ is transversal, i.e. $\boldsymbol{E}'_o = E'_o \mathbf{c}_1, \boldsymbol{H}_t$ has a longitudinal component parallel to $\mathbf{c}_2$, and the wave is not transversal, as asserted.

In the high frequency limit $\omega \to \infty$ of geometrical optics the damping rate p' tends to infinity, the evanescent wave disappears completely, and we have the phenomenon of *total reflection*.

4.8.3 Phase shift.

It remains to enforce the continuous matching of the amplitudes of the incident, reflected and refracted waves at the interface between the two dielectrics. We consider for simplicity the particular case of normal incidence

$$\theta_i = 0 \ , \ \theta_r = \pi \ , \ \boldsymbol{k} = \mathbf{c}_3 = - \, \boldsymbol{k}''$$

and we suppose, without loss of generality, that

$$\boldsymbol{E}_{\rm o} = E_o \mathbf{c}_1 \quad , \quad \mathbf{H}_{\rm o} = H_o \mathbf{c}_2$$

Snell's law (4.128) yields then $\theta_t = 0, \boldsymbol{k}' = \boldsymbol{k} = \mathbf{c}_3$, so that the incident, reflected and transmitted wave depend only on (x_3, t). The continuous matching of $\boldsymbol{n} \wedge \boldsymbol{E}$, $\boldsymbol{n} \wedge \boldsymbol{H}$ at $x_3 = 0$ implies that we must have

$$\boldsymbol{E}'_{\rm o} = E'_o \mathbf{c}_1 \ , \ \boldsymbol{H}'_{\rm o} = H'_o \mathbf{c}_2 \, , \ \boldsymbol{E}''_o \ = E''_o \mathbf{c}_1 \ , \ \boldsymbol{H}''_o \ = H''_o \mathbf{c}_2$$

$$E_o + E''_o \ = E'_o \ , \qquad H_o + H''_o \ = H'_o \tag{4.132}$$

The magnetic amplitudes can be eliminated using eqs. (4.124), (4.125) and (4.127), written for $\boldsymbol{k}' = \boldsymbol{k} = \mathbf{c}_3$ and $\boldsymbol{k} \wedge \mathbf{c}_1 = \mathbf{c}_2$

$$H_o = n_i \sqrt{\frac{\epsilon_o}{\mu_o}} E_o \ , \quad H'_o = n_r \sqrt{\frac{\epsilon_o}{\mu_o}} E'_o \ , \ \ H''_o = -n_i \sqrt{\frac{\epsilon_o}{\mu_o}} E''_o \tag{4.133}$$

In this way eqs. (4.132) become

$$E'_o - E''_o \ = E_o \quad , \quad n_r E'_o + n_i E''_o \ = n_i E_o$$

For given $\boldsymbol{E}_o$ this is a linear algebraic system for the reflected and transmitted electric amplitudes whose unique solution is

$$E''_o = \frac{n_i - n_r}{n_i + n_r} E_o \quad , \qquad E'_o = \frac{2n_i}{n_i + n_r}\, E_o \tag{4.134}$$

Substituting (4.134) into (4.133) yields the reflected and transmitted magnetic amplitudes as

$$H''_o = \frac{n_r - n_i}{n_i + n_r} H_o \ , \ \ H'_o = \frac{2n_r}{n_i + n_r} H_o \tag{4.135}$$

The solution (4.134), (4.135) has the following properties:

(*i*) E''_o, H''_o cannot be taken equal to zero if $n_i \neq n_r$, and therefore the reflected wave cannot disappear.

(*ii*) The Poynting power flux densities $\boldsymbol{E}\wedge\boldsymbol{H}\cdot\boldsymbol{n}$ for the incident, reflected and transmitted waves are [14]

$$P_i := \boldsymbol{E}_i \wedge \boldsymbol{H}_i \cdot \boldsymbol{k} = \mathbf{c}_1 \wedge \mathbf{c}_2 \cdot \mathbf{c}_3 E_o H_o\, cos^2\varphi \equiv n_i \sqrt{\frac{\epsilon_o}{\mu_o}} E_o^2 cos^2\varphi$$

$$P_r := \boldsymbol{E}_r \wedge \boldsymbol{H}_r \cdot \boldsymbol{k}'' = -\mathbf{c}_1 \wedge \mathbf{c}_2 \cdot \mathbf{c}_3 E''_o H''_o\, cos^2\varphi \equiv \left[\frac{n_i - n_r}{n_i + n_r}\right]^2 n_i \sqrt{\frac{\epsilon_o}{\mu_o}} E_o^2\, cos^2\varphi$$

$$P_t := \boldsymbol{E}_t \wedge \boldsymbol{H}_t \cdot \boldsymbol{k}' = \mathbf{c}_1 \wedge \mathbf{c}_2 \cdot \mathbf{c}_3 E'_o H'_o\, cos^2\varphi \equiv \frac{4n_i n_r}{(n_i + n_r)^2}\, n_i \sqrt{\frac{\epsilon_o}{\mu_o}} E_o^2\, cos^2\varphi$$

where $\varphi := \omega\,(t - \frac{\boldsymbol{k}\cdot\boldsymbol{x}}{c_1})$ is the phase. Note that

$$P_r + P_t = P_i.$$

The fact that $P_r > 0$ confirms that the reflected wave travels in the negative x_3-direction, i.e. that $\boldsymbol{k}'' = -\mathbf{c}_3,\ \theta_r = \pi - \theta_i$.

(*iii*) There is a $\pi-$phase shift of the electric or magnetic field upon reflection. Precisely, if $n_i < n_r$ the reflected electric field $\boldsymbol{E}_r$ changes sign with respect to $\boldsymbol{E}_i$, if $n_i > n_r$ this happens to $\boldsymbol{H}_r$ with respect to $\boldsymbol{H}_i$.

(*iv*) $\boldsymbol{D}\cdot\mathbf{c}_3$ and $\boldsymbol{B}\cdot\mathbf{c}_3$ are identically zero for all $\boldsymbol{x}$, and so $\boldsymbol{D}\cdot\boldsymbol{n}$ and $\boldsymbol{B}\cdot\boldsymbol{n}$ match continuously at the interface $x_3 = 0$. It follows that no electric surface charge arises at the interface due to the reflection/refraction process.

[14] since the Poynting vector involves a nonlinear operation (cross product) of $\boldsymbol{E}$ and $\boldsymbol{H}$, the correct expression $Re\boldsymbol{E}\wedge Re\boldsymbol{H}\cdot\boldsymbol{n}$ does not coincide with $Re(\boldsymbol{E}\wedge\boldsymbol{H})\cdot\boldsymbol{n}$

(v) The results for normal incidence are independent of the polarization of the (transversal) incident wave. For oblique incidence this is no longer true, but the continuous matching for $\boldsymbol{D}\cdot\boldsymbol{n}$ and $\boldsymbol{B}\cdot\boldsymbol{n}$ at the interface $x_3 = 0$ follows automatically from that for $\boldsymbol{E}$ and $\boldsymbol{H}$ [35].

4.9 Interference in thin films. Reflection reduction

Consider the light reflected and refracted from a thin dielectric film when a plane linearly polarized wave is normally incident on it. Suppose the given incoming transverse electromagnetic wave is traveling along the $z-$axis in a dielectric medium of permittivity ϵ_1. Part of the wave will be reflected at the first interface $z = 0$ and part transmitted. This transmitted wave will be further reflected internally at the second interface $z = a$ and part transmitted in the dielectric beyond the film. When the internal reflected wave impinges on the first surface, part of it will be will be transmitted and part internally reflected, a.s.o. . These multiple reflections may give rise to interference effects that weaken or cancel altogether the resulting reflected wave in the first medium, contrary to what happens for a single interface.

With this problem in mind, we consider the propagation of an arbitrary plane electromagnetic wave through a layered medium consisting of a semi-infinite dielectric layer $z < 0$ with permittivity ϵ_1, a dielectric layer (slab) $0 < z < a$ with permittivity ϵ_2, and another semi-infinite dielectric layer $z > a$ with permittivity ϵ_3. (In this section we use the notations $x = x_1$, $y = x_2$, $z = x_3$.) This setting encompasses light propagation through a thin film in air (if $\epsilon_1 = \epsilon_3$) or the reflection of light from a coated surface (if $\epsilon_1 < \epsilon_3$).

As in §4.8.3, we assume that:

(i) the wave travels along the $z-$axis (normal incidence);

(ii) the wave is transversal and linearly polarized, with the electric field parallel to the $x-$axis and the magnetic field to the $y-$axis;

(iii) the constant permittivities satisfy $\varepsilon_1 \neq \varepsilon_2, \varepsilon_2 \neq \varepsilon_3$ (otherwise the problem is trivial or is of the type considered in §4.8);

(iv) the magnetic permeability $\mu = \mu_o$ is constant everywhere.

The wave, however, is not assumed to be monochromatic and the problem will be formulated as a boundary value problem for the Maxwell equations and solved using an entirely general method, which applies to any traveling plane waves and can be extended to propagation problems in an arbitrary number of layers and to nonlinear problems [15].

We start from the Maxwell equations (4.1), (4.2) with $\boldsymbol{J} = \rho = 0$. Putting

$$\boldsymbol{E} = E(z,t)\mathbf{c}_1,\ \boldsymbol{H} = H(z,t)\mathbf{c}_2$$

we have $div\boldsymbol{E} = div\boldsymbol{H} = 0$ and the remaining eqs. (4.1), (4.2) reduce to the 2×2 hyperbolic system [16]

$$\frac{\partial H}{\partial z} + \epsilon\frac{\partial E}{\partial t} = 0\ ,\quad \frac{\partial E}{\partial z} + \mu\frac{\partial H}{\partial t} = 0 \qquad (z \in \mathbb{R}\,,\, t \in \mathbb{R}) \tag{4.136}$$

with piecewise constant permittivity

$$\epsilon := \begin{cases} \epsilon_1 & a_o < z < a_1 \qquad a_o = -\infty \\ \epsilon_2 & a_1 < z < a_2 \qquad a_1 = 0,\, a_2 = a \\ \epsilon_3 & a_2 < z < a_3 \qquad a_3 = +\infty \end{cases}$$

and constant permeability $\mu = \mu_o$. For reasons of homogeneity of dimensions it is convenient to introduce the normalized variables

$$\tau = \mathrm{c}_2 t \equiv \frac{t}{\sqrt{\varepsilon_2\mu_o}}\ ,\quad u(z,\tau) = \varepsilon_2 c_2 E(z,t) \equiv \sqrt{\frac{\varepsilon_2}{\mu_o}}E(z,t)\ ,\quad v(z,\tau) = H(z,t)$$

($\mathrm{c}_2 = (\varepsilon_2\mu_o)^{-1/2}$) and the positive quantities

$$h_i := \sqrt{\frac{\varepsilon_i}{\varepsilon_2}} \equiv \frac{c_2}{c_i}\quad ,\qquad i = 1,2,3$$

so that

$$\tau \pm h_i z = (t \pm \frac{z}{c_i})c_2 \quad ,\ i = 1,2,3 \tag{4.137}$$

---

[15] for recent results about the general reflection-transmission problem see e.g. [10] and [14]

[16] this system is symmetrizable and is hyperbolic both in the $z-$variable and in the $t-$variable

Eq. (4.136) becomes then

$$(4.138) \qquad \begin{cases} u_z + v_\tau &= 0 \\ v_z + h_i^2 u_\tau &= 0 \end{cases} \qquad (\, a_{i-1} < x < a_i \, , \, \tau \in \mathbb{R})$$

($i = 1, 2, 3$). The general solution (u_i, v_i) of (4.138) in the i-th layer is given by the superpositions of two arbitrary traveling waves U_i, V_i

$$u_i(z,\tau) = \frac{1}{h_i}\big[U_i(\tau - h_i z) + V_i(\tau + h_i z)\big]\,, \quad v_i(z,\tau) = U_i(\tau - h_i z) - V_i(\tau + h_i z)$$

in each region $a_{i-1} < x < a_i \, , \, \tau \in \mathbb{R}$ ($i = 1, 2, 3$) (Exercise 14). In the theory of partial differential equations U_i, V_i are called Riemann invariants or simple waves.

Eq. (4.137) shows that U_i are waves travelling to the right, V_i are waves traveling to the left,

$$\mathcal{I}(\tau) := U_1(\tau) \quad , \quad \tau \in \mathbb{R}$$

is the given incoming wave,

$$\mathcal{R}(\tau) := V_1(\tau) \quad , \quad \tau \in \mathbb{R}$$

is the (unknown) reflected wave in the first region $z < 0$, and

$$U_3(\tau) := \mathcal{T}(\tau) \, , \quad \tau \in \mathbb{R}$$

is the (unknown) transmitted wave in the third region $z > a$. We tacitly assume that $V_3 \equiv 0$, hence U_3 will be the only wave left for $z > a$. In what follows we will keep the notation (u, v) (and U, V) for the field (u_2, v_2) (and U_2, V_2) inside the slab, where $h_2 = 1$. The general solution of (4.138) can then be written as

(4.139)

$$\begin{aligned} u_1(z,\tau) &= \frac{1}{h_1}\big[\mathcal{I}(\tau - h_1 z) + \mathcal{R}(\tau + h_1 z)\big] \ , \ v_1(z,\tau) = \mathcal{I}(\tau - h_1 z) - \mathcal{R}(\tau + h_1 z) \\ u(z,\tau) &= U(\tau - z) + V(\tau + z) \ , \ v(z,\tau) = U(\tau - z) - V(\tau + z) \\ u_3(z,\tau) &= \frac{1}{h_3}\mathcal{T}(\tau - h_3 z) \quad , \quad v_3(z,\tau) = \mathcal{T}(\tau - h_3 z) \end{aligned}$$

and the wave propagation problem can be stated in the form of the following

Transmission problem: For an arbitrary bounded incident wave $\mathcal{I}(\tau)$ find a bounded solution (4.139) of (4.138) for $(z,\tau)\in\mathbb{R}^2$, such that (u,v) match continuously at the interfaces $z=a_1\equiv 0$, $z=a_2\equiv a$.

The boundedness assumption will be justified later on. The solution of the transmission problem is given by (4.139) provided the unknown functions U,V can be determined uniquely as bounded functions on $\mathbb{R}^2$ by the continuous matching at $z=0$, a. Enforcing the continuous matching at $z=0$ yields

(4.140)
$$u(0,\tau)=u_1(0,\tau)\equiv\frac{1}{h_1}(\mathcal{I}(\tau)+\mathcal{R}(\tau))\ \ ,\ \ v(0,\tau)=v_1(0,\tau)\equiv\mathcal{I}(\tau)-\mathcal{R}(\tau)$$

By combining these equations we immediately see that the unknown field u,v inside the slab must satisfy the boundary condition at $z=0$

(4.141)
$$h_1u(0,\tau)+v(0,\tau)=2\mathcal{I}(\tau)\quad,\qquad\tau\in\mathbb{R}$$

Similarly, enforcing the continuous matching at $z=a$ yields

(4.142)
$$u(a,\tau)=u_3(a,\tau)\equiv\frac{1}{h_3}\mathcal{T}(\tau-h_3a)\ \ ,\quad v(a,\tau)=v_3(a,\tau)\equiv\mathcal{T}(\tau-h_3a)$$

and therefore u, v must satisfy the boundary condition at $z=a$

(4.143)
$$h_3u(a,\tau)-v(a,\tau)=0\quad,\qquad\tau\in\mathbb{R}$$

On the other hand, eqs. (4.140), (4.142) also show that

(4.144)
$$\mathcal{R}(\tau)\ =\ \mathcal{I}(\tau)-v(0,\tau)\ \ ,\quad\mathcal{T}(\tau)=v(a,\tau+h_3a)$$

so that the reflected and transmitted wave are known as soon as the solution $u(z,\tau)$, $v(z,\tau)$ inside the slab $D_a=\{0<z<a,\tau\in\mathbb{R}\}$ is known. Therefore the transmission problem is equivalent to a pure boundary value problem, without initial conditions.

BVP:For an arbitrary bounded function $\mathcal{I}(\tau)$, $\tau\in\mathbb{R}$, find a **bounded** solution (u,v) of the 2×2 hyperbolic system in the slab $D_a=\{0<z<a,\tau\in\mathbb{R}\}$

(4.145)
$$\begin{cases}u_z+\ v_\tau\ =\ 0\\ v_z+u_\tau\ =\ 0\end{cases}$$

satisfying the **impedance boundary conditions** (4.141), (4.143) at $z = 0$ and $z = a$.

(See Exercise 15.) We will consider weak solutions, in the sense of §4.6, so that the functions U, V in (4.139) may be discontinuous as a consequence of discontinuities in $\mathcal{I}(\tau)$. In fact, the solution will be seen to have the same regularity as $\mathcal{I}(\tau)$. If $\mathcal{I}(\tau)$ is continuous and has a continuous bounded first derivative, we will obtain a smooth classical solution, piecewise$-C^1$ in D_a.

Of particular interest is the case where $\mathcal{I}(\tau)$ is periodic with period $\lambda = 2\pi/\omega$, for instance $\mathcal{I}(\tau) = cos\,\omega\tau$ (a monochromatic wave as in §4.8) or $\mathcal{I}(\tau) = sgn(sin\,\omega\tau)$, a periodic train of "square waves". In this case we expect that the solution be also periodic, with the same period. Note that since $\tau = c_2 t$ is actually a space variable, the period λ coincides with the wavelength in the slab.

Proposition 4.9.1 (*i*) *The boundary value problem BVP (and hence the transmission problem) has a unique bounded solution for any bounded* $\mathcal{I}(\tau)$.

(*ii*) *If* $\mathcal{I}(\tau)$ *is periodic, the solution* $u(z,\tau), v(z,\tau)$ *as well as the reflected and transmitted waves* $\mathcal{R}(\tau)$ *and* $\mathcal{T}(\tau)$, *are periodic in* τ *with the same period.*

(*iii*) *The functions* $U(\tau)$, $V(\tau)$, $\mathcal{R}(\tau)$ *and* $\mathcal{T}(\tau)$ *have the same regularity as* $\mathcal{I}(\tau)$.

Proof. The unknowns are (U, V), and from (4.139), (4.141) and (4.143) we have

$$(h_1+1)U(\tau)+(h_1-1)V(\tau) = 2\mathcal{I}(\tau)\,, \quad (h_3-1)U(\tau-a)+(h_3+1)V(\tau+a) = 0$$

that is,

$$U(\tau) - r_1 V(\tau) = \frac{2}{1+h_1}\,\mathcal{I}(\tau)\,, \quad r_2\, U(\tau-a) + V(\tau+a) = 0 \qquad (4.146)$$

where r_1 and r_2 are the reflection coefficients at $z = 0$ and $z = a$,respectively:

$$r_1 = \frac{1-h_1}{1+h_1} \equiv \frac{\sqrt{\varepsilon_2} - \sqrt{\varepsilon_1}}{\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}}\,, \quad r_2 = \frac{h_3-1}{h_3+1} \equiv \frac{\sqrt{\varepsilon_3} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_3} + \sqrt{\varepsilon_2}}$$

Since $\varepsilon_1 \neq \varepsilon_2$, $\varepsilon_2 \neq \varepsilon_3$ (see assumption (iii), we have

$$0 < |r_1 r_2| < 1 \qquad (4.147)$$

and we can write

$$U(\tau) = -\frac{1}{r_2} V(\tau + 2a) \tag{4.148}$$

where $V(\tau)$ is a solution of

$$V(\tau + 2a) + r_1 r_2 V(\tau) = F(\tau) := -\frac{2r_2}{1 + h_1}\mathcal{I}(\tau) \tag{4.149}$$

for $\tau \in \mathbb{R}$. As $|r_1 r_2| < 1$, this difference equation has a unique bounded solution for any bounded F, given by

$$V(\tau) = \sum_{n=0}^{\infty} (-r_1 r_2)^n F(\tau - 2a - 2na) \tag{4.150}$$

(Exercise 16). From eqs. (4.148) and (4.150) we obtain then

$$U(\tau) = -\frac{1}{r_2} V(\tau + 2a) = -\frac{1}{r_2} \sum_{n=0}^{\infty} (-r_1 r_2)^n F(\tau - 2na)$$

so that

$$v(z, \tau) = U(\tau - z) - V(\tau + z) = \sum_{n=0}^{\infty} (-r_1 r_2)^n (-\frac{1}{r_2} F(\tau - 2na - z) - F(\tau - 2a - 2na + z))$$

and the reflected wave follows from eq. (4.144) and manipulations

$$\mathcal{R}(\tau) = \mathcal{I}(\tau) - v(0, \tau) \tag{4.151}$$

$$= -r_1 \mathcal{I}(\tau) - r_2 (1 - r_1^2) \sum_{n=0}^{\infty} (-r_1 r_2)^n \mathcal{I}(\tau - 2a - 2na)$$

The explicit expression for $u(z, \tau) = U(\tau - z) + V(\tau + z)$ and for the transmitted wave

$$\mathcal{T}(\tau) = v(a, \tau + h_3 a)$$

can be immediately derived from the above series expansions for U, V, v.

It is easy to see that $U(\tau), V(\tau), \mathcal{R}(\tau)$ and $\mathcal{T}(\tau)$ are bounded for $\tau \in \mathbb{R}$, have the same regularity as $\mathcal{I}(\tau)$ and, if $\mathcal{I}(\tau)$ is periodic, they are also periodic with the same period.

Further properties of the solution are given as Exercise 17. In particular, part *(iii)* of Exercise 17 shows that the homogeneous problem has unbounded solutions. This can be used to show that uniqueness requires boundedness, and therefore the boundedness requirement is essential in the formulation of **BVP** (and of the **transmission problem**). If, however, the incoming wave $\mathcal{I}(\tau) \equiv 0$ for $\tau \leq \tau_o$, then the series reduces to a finite number of terms for each fixed τ, the restriction $|r_1 r_2| \neq 1$ and the boundedness assumptions can be dropped, and the solution coincides with that of an initial-boundary value problem.

The solution is written as a superposition of the first reflected wave and of waves arising from multiple reflections at the slab walls. Equation (4.151) enables one to easily solve several reflection reduction problems:under suitable conditions the terms in the series (4.151) add up to zero in the periodic case, and the reflected wave disappears entirely.

Proposition 4.9.2 *(half-wave layer). Suppose that the wavelength in the slab λ satisfies*

$$a = \frac{1}{2} m\lambda \ , \quad m = 1, 2, ...$$

and that $r_1 + r_2 = 0$ (i.e. $\varepsilon_1 = \varepsilon_3$). Then $\mathcal{R}(\tau) \equiv 0$.

Proof. Under the stated assumptions $\mathcal{I}(\tau)$ has period $2a$, and the sum of the series (4.151) is

$$\mathcal{R}(\tau) = (-r_1 - r_2 \frac{1 - r_1^2}{1 + r_1 r_2}) \mathcal{I}(\tau) = -\frac{r_1 + r_2}{1 + r_1 r_2} \mathcal{I}(\tau)$$

Therefore $\mathcal{R}(\tau) \equiv 0$ if

$$r_1 + r_2 = \frac{\sqrt{\varepsilon_2} - \sqrt{\varepsilon_1}}{\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}} + \frac{\sqrt{\varepsilon_3} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_3} + \sqrt{\varepsilon_2}} = 0$$

i.e. if $\varepsilon_1 = \varepsilon_3$.

Proposition 4.9.3 *(quarter-wave layer). Suppose $\mathcal{I}(\tau + \lambda/2) = -\mathcal{I}(\tau)$, and*

$$a = \frac{2m - 1}{4} \lambda \qquad m = 1, 2, ...$$

If $r_1 = r_2$ (i.e. $\varepsilon_2 = \sqrt{\varepsilon_1 \varepsilon_3}$), then $\mathcal{R}(\tau) \equiv 0$.

Proof. In this case $\mathcal{I}(\tau)$ has period $4a$, and the sum of the series (4.151) turns out to be

$$\mathcal{R}(\tau) = \frac{r_2 - r_1}{1 - r_1 r_2}\,\mathcal{I}(\tau)$$

Therefore $\mathcal{R}(\tau) \equiv 0$ if

$$r_1 - r_2 = \frac{\sqrt{\varepsilon_2} - \sqrt{\varepsilon_1}}{\sqrt{\varepsilon_2} + \sqrt{\varepsilon_1}} - \frac{\sqrt{\varepsilon_3} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_3} + \sqrt{\varepsilon_2}} = 0$$

that is if $\varepsilon_2 = \sqrt{\varepsilon_1 \varepsilon_3}$,as asserted.

In practice, since a dielectric is always slightly absorbing, the reflection reduction phenomenon in Proposition 4.9.2. and Proposition 4.9.3 is observable only if the integer m is small, so that the slab thickness a is of the order of few wavelengths.

The condition $\varepsilon_2 = \sqrt{\varepsilon_1 \varepsilon_3}$ in Proposition 4.9.3 is approximately satisfied for example by a soap film ($\sqrt{\varepsilon_2/\varepsilon_1} \cong 1.3$) lying on a glass plate (with $\sqrt{\varepsilon_3/\varepsilon_1} \cong 1.69$). This is the principle underlying the behavior of so-called "invisible" glass made by evaporating a thin transparent film on its surface.

4.10 Wave reflection from a system of plane layers.

The problem considered in the previous section can be generalized to a layered medium consisting of m plane dielectric layers $D_k : a_k < z < a_{k+1}$ ($k = 1, ..., m$), with

$$0 = a_1 < a_2 < ... < a_{m+1} \equiv a$$

lying between two semi-infinite homogeneous dielectrics $z < 0$, $z > a_{m+1}$. All $n = m + 2$ media are assumed homogeneous and non-absorbing, with common magnetic permeability $\mu = \mu_o$ and piecewise constant electrical permittivities given by

$$\varepsilon = \varepsilon(z) := \varepsilon_i \quad \text{for} \quad a_{i-1} < z < a_i \qquad (i = 1, ..., n) \tag{4.152}$$

where $a_o := -\infty$, $a_n := +\infty$, $n = m + 2$. As in the previous section, we suppose that a plane linearly polarized wave, traveling to the right in the first medium ($z < 0$), is incident on the system of m plane layers $D_a : 0 \le z \le a$ at arbitrary (non-grazing) incidence. The angles between the propagation directions in each of the media and the normals to the layers (i.e. the $z-$axis) will be denoted by $\theta_k,\ k = 1, ..., n$; thus θ_1 is the angle of incidence, and the $(y, z)-$plane will be considered as the plane of incidence. As a result of multiple reflections at the boundaries of the layers two waves will exist, a "progressive" and a "regressive" one, in each of the media with the exception of the last one ($z > a$), where we expect to have no regressive wave.

For normal incidence ($\theta_1 = 0$, see §4.10.1) the waves will be of TEM (transverse electromagnetic) type, with field vectors

$$\boldsymbol{E} = E(z,t)\mathbf{c}_1 \ ,\ \boldsymbol{H} = H(z,t)\mathbf{c}_2$$

satisfying the Maxwell equations (4.136), where $\varepsilon = \varepsilon(z)$ is defined by (4.152) and $(z, t) \in \mathbb{R}^2$.

For oblique incidence (see §4.10.2) two types of waves are possible, namely, TE (transverse electric) or TM (transverse magnetic) waves, depending on the variables $(y, z, t) \in \mathbb{R}^3$. For TE waves, the electric and magnetic vectors will be

$$\boldsymbol{E} = E_1(y,z,t)\mathbf{c}_1 \ ,\ \boldsymbol{H} = H_2(y,z,t)\mathbf{c}_2 + H_3(y,z,t)\mathbf{c}_3$$

and the Maxwell equations are

$$(4.153)\quad \frac{\partial H_3}{\partial y} - \frac{\partial H_2}{\partial z} = \epsilon\frac{\partial E_1}{\partial t} \quad , \quad \frac{\partial E_1}{\partial y} = \mu\frac{\partial H_3}{\partial t} \quad , \quad \frac{\partial E_1}{\partial z} + \mu\frac{\partial H_2}{\partial t} = 0$$

For TM waves, the electric and magnetic fields

$$\boldsymbol{E} = E_2(y,z,t)\mathbf{c}_2 + E_3(y,z,t)\mathbf{c}_3 \ ,\ \boldsymbol{H} = H_1(y,z,t)\mathbf{c}_1$$

will satisfy the Maxwell equations

$$(4.154)\qquad \frac{\partial E_3}{\partial y} - \frac{\partial E_2}{\partial z} = -\mu\frac{\partial H_1}{\partial t} \ , \ -\frac{\partial H_1}{\partial y} = \epsilon\frac{\partial E_3}{\partial t} \ , \ \frac{\partial H_1}{\partial z} = \epsilon\frac{\partial E_2}{\partial t}$$

As in §4.9, we look for solutions to the following

Transmission problem: For an arbitrarily given, bounded incident wave in the first region $z < 0$, find a bounded solution $\boldsymbol{E}, \boldsymbol{H}$ satisfying the continuous matching conditions for the tangential components of $(\boldsymbol{E},\boldsymbol{H})$ and the normal components of $(\varepsilon(z)\boldsymbol{E},\boldsymbol{H})$ at all the interfaces $z=a_{\mathrm{j}}$, $j=1,...,m+1$.

This linear transmission problem can be reduced to solving a single linear functional equation for one of the Riemann invariants.

The method of analysis given here for electromagnetic waves enables us to discuss the phenomenon of reflection reduction in optical systems and interference lightfilters, and can easily be adapted to acoustic waves.

4.10.1 Normal incidence.

It is convenient to work with normalized variables

$$\tau=\frac{t}{\sqrt{\varepsilon_2\mu}}\ ,\ \ u(z,\tau)=\sqrt{\frac{\varepsilon_2}{\mu}}E(z,\tau\sqrt{\varepsilon_2\mu})\ ,\ \ v(z,\tau)=H(z,\tau\sqrt{\varepsilon_2\mu})$$

Taking eq. (4.152) into account, the Maxwell equations (4.136) become then

$$u_z+v_\tau=0\ ,\ \ v_z+h_k^2u_\tau=0\quad(a_{k-1}<z<a_k\,,\,\tau\in\mathbb{R})\tag{4.155}$$

where

$$h_k=\sqrt{\frac{\varepsilon_k}{\varepsilon_2}}\qquad(k=1,...,n=m+2)\tag{4.156}$$

The general solution (u_i,v_i) of (4.155) in the $k-$th medium can be written in terms of the Riemann invariants U_k, V_k :

$$\begin{aligned}u_k(z,\tau)&=\frac{1}{h_k}(U_k(\tau-h_kz)+V_k(\tau+h_kz))\\ v_k(z,\tau)&=U_k(\tau-h_kz)-V_k(\tau+h_kz)\qquad\scriptstyle(a_{k-1}<z<a_k\ ,\ \tau\in\mathbb{R})\end{aligned}\tag{4.157}$$

(see Exercise 14). Here $V_1(\tau)=\mathcal{R}(\tau)$ represents the reflected wave for $z<0$ and $U_n(\tau)=\mathcal{T}(\tau)$ the transmitted wave for $z>a$. The continuous matching conditions at $z=a_j$

$$u_{j+1}(a_j,\tau)=u_j(a_j,\tau)\ ,\ \ v_{j+1}(a_j,\tau)=v_j(a_j,\tau)\qquad(j=1,...,m+1)$$

become
(4.158)

$$\begin{aligned}&Z_{j+1}\big[U_{j+1}(\tau-h_{j+1}a_j)+V_{j+1}(\tau+h_{j+1}a_j)\big]\\&=Z_j\big[U_j(\tau-h_ja_j)+V_j(\tau+h_ja_j)\big]\\&U_{j+1}(\tau-h_{j+1}a_j)-V_{j+1}(\tau+h_{j+1}a_j)=U_j(\tau-h_ja_j)-V_j(\tau+h_ja_j)\big]\end{aligned}$$

$(j=1,...,m+1,\ \tau\in\mathbb{R})$, where Z_j are the wave impedances

(4.159)
$$Z_j := \sqrt{\frac{\mu}{\varepsilon_j}}$$

and h_j defined by (4.156) coincides with the ratio of the refractive indices:

$$h_j := \sqrt{\frac{\varepsilon_j}{\varepsilon_2}} \equiv \frac{n_j}{n_2}$$

The Riemann invariants U_k, V_k must therefore be determined as solutions of (4.158) corresponding to a given incident wave in the first medium $z<0$

$$U_1(\tau) = \mathcal{I}(\tau) \qquad (\tau\in\mathbb{R})$$

and to a vanishing regressive wave in the last medium $z>a$

$$V_n(\tau)\equiv 0 \qquad (\tau\in\mathbb{R})$$

As in the case of a single layer, eqs. (4.158) can be reduced by elimination to a single linear functional equation for one of the unknown Riemann invariants, say, $V_2(\tau)$. Under a suitable restriction on the wave impedances Z_i, this equation has a unique bounded solution $V=V_2(\tau)$ for bounded $\mathcal{I}(\tau)$. This in turn yields a unique bounded solution (u,v) (in a generalized sense) to our transmission problem. If $\mathcal{I}(\tau)$ is continuous and has a continuous bounded first derivative, we obtain a smooth classical solution, piecewise$-C^1$ in D_a. If $\mathcal{I}(\tau)$ is periodic, the solution (u,v) is periodic, with the same period.

We proceed to give a detailed analysis for $m=2$ layers, and by induction, to derive a formula for the wave $\mathcal{R}(\tau)$ reflected by an arbitrary multilayered coating with m layers (the case of a single layer has already been examined in §4.9). To this purpose, we define the reflection coefficients

(4.160)
$$r_j := \frac{Z_j-Z_{j+1}}{Z_j+Z_{j+1}} \equiv \frac{\sqrt{\varepsilon_{j+1}}-\sqrt{\varepsilon_j}}{\sqrt{\varepsilon_{j+1}}+\sqrt{\varepsilon_j}} \qquad (j=1,...,m+1)$$

This definition implies that $0 < |r_j| < 1$. For brevity, we also put

$$\delta_j := a_j - a_{j-1} \ , \ \ d_j := 2h_j \ \delta_j \qquad (j = 1, ..., m+1) \tag{4.161}$$

so that here $d_2 = 2a_2$. It is particularly interesting to investigate the possibility of absence of reflection, i.e. to see whether it is possible to have $\mathcal{R}(\tau) \equiv 0$.

(i) Two layers: $m = 2$. From eqs. (4.158), all Riemann invariants U_i , V_i can be expressed in terms of $V = V_2$, which must then satisfy the functional equation

$$\sum_{j=1}^{4} A_j V(\tau + b_j) = F(\tau) \ , \qquad \tau \in \mathbb{R} \tag{4.162}$$

with

$$A_1 = r_1 r_3 \ , \ \ A_2 = r_2 r_3 \ , \ A_3 = r_1 r_2 \ , \ \ A_4 = 1$$
$$b_1 = 0 \ \ , \ \ \ b_2 = d_2 \ , \ \ \ b_3 = d_3 \ \ , \ b_4 = d_2 + d_3$$

and

$$F(\tau) \ = -(1 + r_1)\big[r_3 \mathcal{I}(\tau) + r_2 \mathcal{I}(\tau + d_3)\big] \tag{4.163}$$

(Exercise 18 and 19). Note that by definition, $b_4 > 0$ is the highest increment among the b_j's, and $b_1 = 0$ is the lowest.

Proposition 4.10.1 *Under the sufficient assumption*

$$\sum_{j=1}^{3} |A_j| \equiv |r_1 r_2| + |r_2 r_3| + |r_1 r_3| \ < |A_4| = \ 1 \tag{4.164}$$

eq. (4.162) has a unique bounded solution for bounded $\mathcal{I}(\tau)$, given by the uniformly and absolutely convergent series

$$V(\tau) \ = \sum_{n_1, n_2 n_3 = 0}^{\infty} M_{n_1 n_2 n_3} F(\tau - \Delta_{n_1 n_2 n_3}) \ , \qquad \tau \in \mathbb{R} \tag{4.165}$$

where the coefficients $M_{n_1 n_2 n_3}$ and the delays $\Delta_{n_1 n_2 n_3}$ are given by

$$M_{n_1 n_2 n_3} := \frac{(n_1 + n_2 + n_3)!}{n_1! \, n_2! \, n_3!} (-r_3 r_2)^{n_1} (-r_2 r_1)^{n_2} (-r_1 r_3)^{n_3}$$

$$\Delta_{n_1 n_2 n_3} := (1 + n_2 + n_3)d_2 + (1 + n_1 + n_3)d_3$$

and F is defined by eq. (4.163). Proof. See Theorem 1 in the Appendix.

The reflected wave $V_1(\tau) = \mathcal{R}(\tau)$ is found to be given by

$$\mathcal{R}(\tau) = -r_1 \mathcal{I}(\tau) + (1 - r_1) \sum_{n_1, n_2 n_3 = 0}^{\infty} \mathrm{M}_{n_1 n_2 n_3} F(\tau - \Delta_{n_1 n_2 n_3}) \tag{4.166}$$

(cfr. (4.151)). If $\mathcal{I}(\tau)$ is periodic or antiperiodic with period p or antiperiod $\overline{p}$, respectively:

$$\mathcal{I}(\tau + p) = \mathcal{I}(\tau) \quad \text{or} \quad \mathcal{I}(\tau + \overline{p}) = -\mathcal{I}(\tau)$$

so is the solution $V(\tau)$, and conditions for absence of reflection can be easily determined. We consider here the following typical cases:

(a) If $d_2 = kp,\ d_3 = lp$ (with $k, l = 1, 2, ...$) then

$$\mathcal{R}(\tau) = -\frac{r_1 + r_2 + r_3 + r_1 r_2 r_3}{1 + r_1 r_2 + r_2 r_3 + r_1 r_3} \mathcal{I}(\tau)$$

Hence

$$\mathcal{R}(\tau) \equiv 0 \quad \text{for} \quad r_1 + r_2 + r_3 + r_1 r_2 r_3 = 0, \ \text{i.e. for} \quad \varepsilon_1 = \varepsilon_4$$

(b) If $d_2 = k\overline{p},\ d_3 = l\overline{p}$ (with k, l odd) then

$$\mathcal{R}(\tau) = -\frac{r_1 - r_2 + r_3 - r_1 r_2 r_3}{1 - r_1 r_2 - r_2 r_3 + r_1 r_3} \mathcal{I}(\tau)$$

Hence

$$\mathcal{R}(\tau) \equiv 0 \quad \text{for } r_1 - r_2 + r_3 - r_1 r_2 r_3 = 0\,, \quad \text{i.e. for} \quad \varepsilon_3 \sqrt{\varepsilon_1} = \varepsilon_2 \sqrt{\varepsilon_4} \tag{4.167}$$

(c) If $d_2 = k\overline{p},\ d_3 = l\overline{p}$ (with l odd, k even) then

$$\mathcal{R}(\tau) = -\frac{r_1 + r_2 - r_3 - r_1 r_2 r_3}{1 + r_1 r_2 - r_2 r_3 - r_1 r_3} \mathcal{I}(\tau)$$

Hence

$$\mathcal{R}(\tau) \equiv 0 \quad \text{for} \quad r_1 + r_2 - r_3 - r_1 r_2 r_3 = 0\ , \quad \text{i.e. for} \quad \varepsilon_3 = \sqrt{\varepsilon_1 \varepsilon_4} \tag{4.168}$$

(d) If $d_2 = k\overline{p}$, $d_3 = l\overline{p}$ (with l even, k odd) then

$$\mathcal{R}(\tau) = -\frac{r_1 - r_2 - r_3 + r_1 r_2 r_3}{1 - r_1 r_2 + r_2 r_3 - r_1 r_3}\mathcal{I}(\tau)$$

Hence

$$\mathcal{R}(\tau) \equiv 0 \quad \text{for} \;\; r_1 - r_2 - r_3 + r_1 r_2 r_3 = 0 \;, \quad \text{i.e. for} \;\; \varepsilon_2 = \sqrt{\varepsilon_1 \varepsilon_4} \tag{4.169}$$

(ii) Three layers: $m = 3$. In this case the functional equation for V has eight terms, $A_8 = 1$, $b_8 > 0$ is the highest increment among the b_j's, $b_1 = 0$ is the lowest, and eq. (4.164) must be replaced by

$$\sum_{j=1}^{7} |A_j| \equiv \tag{4.170}$$

$$|r_1 r_2| + |r_2 r_3| + |r_1 r_3| + |r_1 r_4| + |r_2 r_4| + |r_3 r_4| + |r_1 r_2 r_3 r_4| < |A_8|$$

Details are left as an exercise.

(iii) m layers ($m > 1$). By induction one finds the series expansion for the reflected wave

$$\mathcal{R}(\tau) = -\sum_{k=1}^{m+1} r_k \Big[\mathcal{I}(\tau - \zeta_{1,k}) + \sum_{n=1}^{k} r_n \sum_{i=n+1}^{m+1} r_i \mathcal{I}(\tau - \zeta_{1,k} - \zeta_{n+1,i})\Big] + \ldots.. \tag{4.171}$$

where

$$\zeta_{j,k} := \sum_{n=j}^{k} d_n$$

Each term of the series (4.171) is called a multiple wave. The series contains only multiple waves of odd order, since an odd number of reflections must occur in order to contribute to the reflected wave $\mathcal{R}$ in the first medium.

Proposition 4.10.2 *If $\mathcal{I}(\tau)$ has bounded support $[\tau_o, \tau_1]$, then*

(i) *the support of $\mathcal{R}(\tau)$ is $[\tau_o, +\infty)$, and*

(ii) $\mathcal{R}(\tau) \to 0$ *as* $\tau \to +\infty$.

Proof. See Theorem 2 in the Appendix. This result is physically intuitive owing to the decay due to the multiple reflections at the layer interfaces.

The series (4.171) converges under assumptions on the reflection coefficients similar to (4.164) and (4.170), and the terms replaced by dots are usually negligible. If the r_i do not satisfy the convergence assumptions, the series (4.171) can still be interpreted as an asymptotic expansion

$$\mathcal{R}(\tau) = -\sum_{k=1}^{m+1} r_k \Big[\mathcal{I}(\tau - \zeta_{1,k}) + \sum_{n=1}^{k} r_n \sum_{i=n+1}^{m+1} r_i \mathcal{I}(\tau - \zeta_{1,k} - \zeta_{n+1,i})\Big] + O(r^3)$$

as $r \to 0,$ where

$$r := max_i |r_i|$$

$(0 < r < 1)$. Keeping only the contributions due to the primary waves

$$\mathcal{R}(\tau) = -\sum_{k=1}^{m+1} r_k\ \mathcal{I}(\tau - \sum_{j=1}^{k} d_j) \tag{4.172}$$

yields an estimate of the full solution up to terms of order $O(r^2)$, and this approximation is acceptable in many cases (typically, $r^2 \cong 0.05$ or less). If $\mathcal{I}(\tau)$ is periodic or antiperiodic with period p or antiperiod $\overline{p}$, respectively, one can investigate reflection reduction due to a multilayered coating using eq. (4.172).

(*a*) If $d_2 + ... + d_k = n_k p$ (n_k integers, $k = 2, ..., m+1$) then

$$\mathcal{R}(\tau) \equiv O(r^2) \quad \text{for} \quad \sum_{k=1}^{m+1} r_k = 0$$

(*b*) If $d_2 + ... + d_k = n_k \overline{p}$ (n_k integers, $k = 2, ..., m+1$) then

$$\mathcal{R}(\tau) \equiv O(r^2) \quad \text{for} \quad \sum_{k=1}^{m+1} (-1)^{n_k} r_k = 0$$

The advantage of increasing the number m of layers is to broaden the window with low reflectance.

Suppose $\mathcal{I}(\tau)$ is the rectangular pulse of duration T

$$\mathcal{I}(\tau) = \begin{cases} 1 & \text{for } |\tau| \le T \\ 0 & |\tau| > T \end{cases}$$

with $0 < T < \frac{1}{2} sup_j(d_j)$. Then eq. (4.172) shows that the reflected wave satisfies

$$\mathcal{R}(\tau) = -r_j + O(r^2) \quad \text{for} \ \ |\tau - d_2 - ... - d_j| \le T \quad (j = 2, ..., m+1)$$

so that the reflection coefficients r_j, or the permittivities ε_j , can be estimated from $\mathcal{R}(\tau)$ by inspection. Also the $d_j's$ can be reconstructed by varying the value of T. In this way an inverse problem is (approximately) solved.

4.10.2 Oblique incidence.

The solution for TE and TM waves at oblique incidence can be formally reduced to that for normal incidence. We consider the two cases separately.

A. TE waves: $\boldsymbol{E}= E_1(y,z,t)\mathbf{c}_1$,$\boldsymbol{H}= H_2(y,z,t)\mathbf{c}_2+H_3(y,z,t)\mathbf{c}_3$. Introducing the normalized variables

$$u(y,z,\tau) = \sqrt{\frac{\varepsilon_2}{\mu}} E_1(y,z,t) \ \ , \ \ v(y,z,\tau) = H_2(y,z,t) \ , \ \ w(y,z,\tau) = H_3(y,z,t)$$

with $t = \tau\sqrt{\varepsilon_2 \mu}$, the Maxwell equations (4.153) become

$$(4.173) \quad v_z + h_k^2 u_\tau = w_y, \ \ u_z + v_\tau = 0 \ , \ \ u_y = w_\tau \qquad (a_{k-1} < z < a_k \ , \ \tau \in \mathbb{R})$$

where h_k is defined by eq. (4.156), and $k = 1, ..., m+2$. We restrict our attention to plane wave solutions of (4.173) in the $k-$th layer $a_{k-1} < z < a_k$ of the form

(4.174)

$$u_k = h_k^{-1}\big[U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) + V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k)\big]$$
$$v_k = cos\theta_k\big[U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) + V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k)\big]$$
$$w_k = -sin\theta_k\big[U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) + V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k)\big]$$

$(k = 1, ..., m+2)$, with $U_1(\tau) = \mathcal{I}(\tau)$ the given incident wave, $V_1(\tau) = \mathcal{R}(\tau)$ the unknown reflected wave for $z < 0$, $U_n(\tau)$ the unknown transmitted wave for $z > a$,and $V_n(\tau) \equiv 0$. The continuous matching conditions at $x = a_j$

$$u_{j+1}(y,a_j,\tau) = u_j(y,a_j,\tau) \ , \ \ v_{j+1}(y,a_j,\tau) = v_j(y,a_j,\tau) \ , \ w_{j+1}(y,a_j,\tau) = w_j(y,a_j,\tau)$$

(with $j = 1, ..., m+1$ and $(y,\tau) \in \mathbb{R}^2$) yield Snell's refraction law

$$(4.175) \qquad h_k sin\theta_k = h_1 sin\theta_1 \qquad \text{for all} \ \ k = 2, ..., m+2$$

If we define here the reflection coefficients as

$$r_j := \frac{\sqrt{\varepsilon_{j+1}}\, cos\, \theta_{j+1} - \sqrt{\varepsilon_j}\, cos\, \theta_j}{\sqrt{\varepsilon_{j+1}}\, cos\, \theta_{j+1} + \sqrt{\varepsilon_j}\, cos\, \theta_j} \qquad (j = 1, ..., m+1)$$

the discussion and the results are formally the same as for normal incidence, at least if

$$h_1 |sin\theta_1| < h_k \quad \text{for all } k = 1, ..., m+2 \tag{4.176}$$

By force of Snell's law, this assumption implies $|cos\, \theta_k| > 0$ for all k, so that total reflection and evanescent waves are ruled out at all the layer interfaces $z = a_k$. If, on the contrary, assumption (4.176) is not satisfied, complex quantities arise, according to the recipe of §4.8.2, and the analysis requires some care. Suppose for example that j is the smallest integer such that $h_1|sin\theta_1| = h_{j+1}$, so that $cos\, \theta_{j+1} = 0$: then $r_j = -1$ and all the subsequent reflection coefficient can be taken as zero ($r_k := 0$ for $k > j$), since the layers with $k > j$ have no effect on the solution. If, on the other hand, $h_1|sin\theta_1| > h_{j+1}$ then

$$\theta_{j+1} = \frac{\pi}{2} + i\beta \;\; \Rightarrow \;\; sin\, \theta_{j+1} = cosh\, \beta\, , \;\; cos\, \theta_{j+1} = -isinh\, \beta$$

so that some of the U_k and V_k become complex quantities.

B. TM waves: $\boldsymbol{E} = E_2(y, z, t)\mathbf{c}_2 + E_3(y, z, t)\mathbf{c}_3$, $\boldsymbol{H} = H_1(y, z, t)\mathbf{c}_1$. Introducing the normalized variables

$$u(y, z, \tau) = \sqrt{\frac{\varepsilon_2}{\mu}} E_2(y, z, t) \;\; , \; v(y, z, \tau) = H_1(y, z, t), \;\; w(y, z, \tau) = \sqrt{\frac{\varepsilon_2}{\mu}} E_3(y, z, t)$$

with $t = \tau\sqrt{\varepsilon_2 \mu}$, the Maxwell equations (4.154) become

$$v_z - h_k^2 u_\tau = 0\, , \;\; u_z - v_\tau = w_y \;\, , \;\; v_y + h_k^2 w_\tau = 0 \qquad (a_{k-1} < z < a_k\, , \; \tau \in \mathbb{R}) \tag{4.177}$$

where h_k is defined by eq. (4.156), and $k = 1, ..., m+2$. Again, we restrict attention to plane wave solutions of (4.177) in the $k-$th layer $a_{k-1} < z < a_k$ of the form

$$\begin{aligned}
u_k &= -h_k^{-1} cos\theta_k \big[U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) + V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k)\big] \\
v_k &= \;\, U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) - V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k) \\
w_k &= h_k^{-1} sin\theta_k \big[U_k(\tau - h_k z\, cos\theta_k - h_k y\, sin\theta_k) - V_k(\tau + h_k z\, cos\theta_k - h_k y\, sin\theta_k)\big]
\end{aligned}$$

($k = 1, ..., m+2$). The continuous matching conditions yield again Snell's law (4.175). If we define here the reflection coefficients as

$$r_j := \frac{\sqrt{\varepsilon_{j+1}}\, cos\, \theta_j \; - \sqrt{\varepsilon_j}\, cos\, \theta_{j+1}}{\sqrt{\varepsilon_{j+1}}\, cos\, \theta_j + \sqrt{\varepsilon_j}\, cos\, \theta_{j+1}} \qquad (j = 1, ..., m+1)$$

the discussion is similar to that carried above for TE waves.

APPENDIX to section 10 : THE FUNCTIONAL EQUATION

The transmission problem considered in §4.10 requires to study existence and uniqueness of bounded solutions, for a given bounded source term $F(\tau)$, of the functional equation

$$\sum_{\alpha=0}^{N} A_\alpha V(\tau + b_\alpha) = F(\tau) \qquad (\tau \in \mathbb{R}) \tag{4.178}$$

where $A_1, ..., A_N$ are arbitrary real coefficients, and we assume without loss of generality that

$$0 = b_o < b_1 < ... < b_N$$

There are no initial data, and the solution must be found for all $\tau \in \mathbb{R}$.

Since (4.178) is linear, the solution will be unique if and only if the homogeneous equation

$$\sum_{\alpha=0}^{N} A_\alpha V_o(\tau + b_\alpha) = 0 \qquad (\tau \in \mathbb{R}) \tag{4.179}$$

has only the trivial solution $V_o\,(\tau\,) \equiv 0.$ If the increments b_α are commensurable:

$$b_\alpha = n_\alpha h \ , \ \ n_\alpha \text{ integers}, \quad 0 = n_0 < n_1 < ... < n_N\,, \quad h > 0$$

then (4.179) reduces to the difference equation with constant coefficients

$$\sum_{k=0}^{K} A'_k V_o(\tau + kh) = 0 \tag{4.180}$$

where

$$A'_0 = A_0 \quad , \quad A'_k = \begin{cases} 0 & \text{if } \ k \neq n_j \\ A_j & \text{if } \ k = n_j \end{cases} \tag{4.181}$$

and $K \geq N$ is a suitable integer whose value follows from (4.181). A well-known classical result says that the general solution of (4.180) is given by

$$V_o(\tau) = Re\Big\{ \sum_{k=0}^{K} w_k(\tau) z_k^{\tau/h} \Big\} \tag{4.182}$$

where the z_k's are the (real or complex) roots of multiplicity m_k of the characteristic polynomial

$$P(z) = \sum_{k=0}^{K} A'_k z^k$$

and

$$w_k(\tau) = \sum_{i=0}^{m_k - 1} p_{ik}(\tau) \tau^i$$

are polynomials of degree $m_k - 1$ with arbitrary coefficients $p_{ik}(\tau)$ periodic of period h. If there is a root z_l of unit modulus, then taking all $p_{ik}(\tau) \equiv 0$ for $i \neq 0$ and $k \neq l$ yields a bounded solution $V_o(\tau)$ of the homogeneous equation. Conversely, if (4.179) has a bounded solution (4.182), then necessarily there must be at least one root z_k with $|z_k| = 1$. In other words, if the b_α's are commensurable the necessary and sufficient condition for uniqueness of solutions to (4.178) is that

$$|z_k| \neq 1 \ \text{ for all } k$$

If this condition is satisfied, the unique bounded solution can be constructed in the form of a power series in the z_k.

In the case of generic increments b_α, not necessarily commensurable, a sufficient condition for existence and uniqueness is strict dominance of one of the coefficients of the equation. By rescaling, we can always assume that

$$|A_\alpha| \leq 1 \qquad \text{for all } \ \alpha = 0, ..., N$$

Theorem 4.10.3 *Suppose that*

$$\sum_{\alpha\neq\beta}^{0...N} |A_\alpha| < |A_\beta| = 1 \tag{4.183}$$

for some integer β, $0 \leq \beta \leq N$. Then, for any bounded $F(\tau)$, eq. (4.178) has a unique bounded solution $V(\tau)$, which depends continuously on $F(\tau)$ and has the same smoothness properties as $F(\tau)$. Furthermore if $F(\tau)$ is periodic $V(\tau)$ is also periodic, with the same period.

Proof. Let V_o satisfy (4.179). Then for every $\tau_o \in \mathbb{R}$ and every integer q we have

$$|V_o(\tau_o)| < \left[\sum_{\alpha\neq\beta}^{0...N} |A_\alpha| \right]^q \sup_{\tau\in\mathbb{R}} |V_o(\tau)|$$

so that, if (4.183) is satisfied and $V_o(\tau)$ is bounded on $\mathbb{R}$, $V_o(\tau) \equiv 0$. This proves uniqueness. The unique bounded solution is given by the uniformly and absolutely convergent series

(4.184)

$$V(\tau) = \sum_{k_1,...,k_N}^{0...\infty} \frac{(k_1 + ... + k_N)!}{k_1!...k_N!} (-A_{m_1})^{k_1} \; ....(-A_{m_N})^{kn} F(\tau - \Delta_{k_1...k_N})$$

where $(m_1...m_N)$ denotes an arbitrary permutation of the $N-1$ integers $0 \leq \alpha \neq \beta \leq N$, and

$$\Delta_{k_1...k_N} := b_\beta + k_1(b_\beta - b_{m_1}) + ... + (b_\beta - b_{m_N})$$

Eq. (4.184) implies the estimate

$$\sup_{\tau\in\mathbb{R}} |V(\tau)| \leq M_\beta \sup_{\tau\in\mathbb{R}} |F(\tau)| \;,\; M_\beta = \left[1 - \sum_{\alpha\neq\beta}^{0...N} |A_\alpha| \right]^{-1}$$

which proves boundedness and continuous dependence. All the other assertions are obvious.

In the applications to the wave problem of §4.10, the index $j = \alpha + 1$ goes from 1 to $J = N + 1$ and (4.183) is satisfied with $\beta = N$, that is,

$$|A_j| < |A_J| = 1$$

for all $j = 1, ...,$J-1. For example, in the case of eq. (4.162) for two layers, $J = 4$, $\beta = 3$ and, by a suitable permutation (m_1, m_2, m_3), eq. (4.184) coincides with (4.165). The fact that (4.183) holds with $\beta = N$ has an important consequence which may be viewed as a sort of causality principle.

Theorem 4.10.4 *If $|A_\alpha| < |A_N| = 1$ for all $\alpha = 0, ..., N-1$, and if $F(\tau)$ has bounded support $[\tau_o, \tau_1]$, then the support of $V(\tau)$ is $[t_o, +\infty)$ and $V(\tau) \to 0$ as $\tau \to +\infty$.*

The proof of this theorem follows by inspection.

Remark 15. The homogeneous equation (4.179), even under assumption (4.183), has nontrivial unbounded solutions (Exercise 19). If (4.183) is not satisfied, (4.179) has non-trivial bounded solutions (Exercise 20). Therefore the boundedness assumption for $V(\tau)$ and $F(\tau)$ as well as assumption (4.183) cannot be dispensed with.

Exercises

Exercise 1. Hint: Use the identities

$$div(f\boldsymbol{J}) \equiv f div\, \boldsymbol{J} + \boldsymbol{J} \cdot grad\, f \ , \ \ curl(f\boldsymbol{v}) \equiv f curl(\boldsymbol{v}) - grad\, f \wedge \boldsymbol{v}$$

Exercise 2. Find the phase and group velocities for the evanescent wave, and show that $v_f v_g = c^2$.

Answer:

$$v_f = \frac{\omega}{p} = \frac{c}{\sqrt{1 + |\mathbf{p}'|^2 c^2/\omega^2}} \ , \ v_g = \frac{d\omega}{dp} = c\sqrt{1 + |\mathbf{p}'|^2 c^2/\omega^2} \tag{(E1)}$$

whence $v_f v_g = c^2$. The refractive index is

$$n_r := \frac{c_o}{v_f} = \frac{c_o}{c}\sqrt{1 + |\mathbf{p}'|^2 c^2/\,\omega^2} \tag{(E2)}$$

Therefore if $|\boldsymbol{p}'| \propto \omega$, as in the case of total reflection in optics (§4.8.2), n_r is independent of ω.

Exercise 3 (Gaussian beams). The wave equation

$$(*) \qquad u_{tt} = c^2 \left(u_{xx} + u_{yy} + u_{zz}\right)$$

has special solutions of the form

$$u = w(x, y, \tau)e^{ih(z+ct)} \quad , \quad \tau := z - ct$$

for every $h > 0$, provided w satisfies the two-dimensional Schrödinger equation

$$ihw_\tau = -\frac{1}{4}(w_{xx} + w_{yy})$$

One solution is the Gaussian beam

$$(**) \qquad w = \frac{1}{\alpha + i\tau} exp\left[-\frac{h(x^2 + y^2)}{\alpha + i\tau}\right]$$

parametrized by the complex number $\alpha \neq -i\tau$. By superposition we find solutions of (*) of the form

$$u(x, y, z, t) = \int_0^\infty \frac{e^{ih(z+ct)}e^{-\frac{h(x^2+y^2)}{\alpha(h)+i(z-ct)}}}{\alpha(h) \, + \, i(z - ct)} \, dh$$

for any bounded $\alpha(h) > 0$ (say). These special solutions are of interest in optics.

Taking $h = 0$ shows that eq. (*) admits traveling wave solutions in z of the form

$$u = w(x, y, z - ct)$$

provided $w_{xx} + w_{yy} = 0$ for every z and t. Solutions of this kind cannot be bounded unless they are independent of x and y.

Indeed, for $\eta := z + ct, \xi := z - ct$ and

$$u = w(x, y, \xi)e^{ih\eta}$$

we have $u_{\eta\xi} = ihu_\xi = ihw_\xi e^{ih\eta}$, and

$$u_{xx} + u_{yy} + u_{zz} - c^2 u_{tt} = u_{xx} + u_{yy} + 4u_{\eta\xi} = e^{ih\eta}(w_{xx} + w_{yy} + 4ihw_\xi)$$

So u satisfies eq. (*) if w satifies the above Schrödinger equation. (**) is the fundamental solution of the Schrödinger equation, parametrized by

α. Finally Liouville's theorem for harmonic functions says that w does not depend on x, y if it is harmonic and bounded [2].

Exercise 4. By applying the Gauss Lemma we have

$$M_r = \frac{1}{4\pi}\int_\Omega \frac{\partial}{\partial r} f(\boldsymbol{x}+r\boldsymbol{\nu})\,d\Omega = \frac{1}{4\pi}\int_\Omega \boldsymbol{\nu}\cdot grad_x f(\boldsymbol{x}+r\boldsymbol{\nu})\,d\Omega \quad \Rightarrow$$
$$\Rightarrow \quad \lim_{r\to 0} M_r = \frac{1}{4\pi} grad_x f(\boldsymbol{x})\cdot\int_\Omega \boldsymbol{\nu}\,d\Omega \;=\; \frac{1}{4\pi} grad\, f\cdot\int_{|\boldsymbol{y}|\le 1} grad(1)\,d\boldsymbol{y} = 0$$

whence $M \in C^1(\mathbb{R}^4)$. But, since $f \in C^2(\mathbb{R}^3)$, we can also write

$$\begin{aligned} M_r &= \frac{1}{4\pi r^2}\int_{\Sigma_r} \boldsymbol{\nu}\cdot grad_y f(\boldsymbol{y})\,dS = \frac{1}{4\pi r^2}\int_{B_r} \Delta_3 f(\boldsymbol{y})\,d\boldsymbol{y} \\ &= \frac{1}{4\pi r^2}\int_0^r d\rho\rho^2 \int_\Omega \Delta_3 f(\boldsymbol{x}+\rho\boldsymbol{\nu})\,d\Omega \end{aligned}$$

where B_r is the ball $|\boldsymbol{y}-\boldsymbol{x}| \le r$, and moreover $\Delta_3 M f = M\Delta_3 f$. Hence

$$\begin{aligned} M_{rr} &= \frac{1}{4\pi}\int_\Omega \Delta_3 f(\boldsymbol{x}+r\boldsymbol{\nu})\,d\Omega - \frac{2}{4\pi r^3}\int_0^r d\rho\rho^2\int_\Omega \Delta_3 f(\boldsymbol{x}+\rho\boldsymbol{\nu})\,d\Omega \\ &= M\Delta_3 f - \frac{2}{r}M_r = \Delta_3 M - \frac{2}{r}M_r \end{aligned}$$

so that M satisfies the Darboux equation. Finally, letting $r \to 0$ yields

$$M_{rr} \;\to\; \Delta_3 f(\boldsymbol{x}) - \Delta_3 f(\boldsymbol{x})\frac{2}{4\pi r^3}\frac{4\pi r^3}{3} = \frac{1}{3}\Delta_3 f(\boldsymbol{x})$$

and since also $Mf\{\boldsymbol{x}, r\} \to f(\boldsymbol{x})$ and $M_r f\{\boldsymbol{x}, r\} \to 0$ as $r \to 0$, we have $M \in C^2(\mathbb{R}^4)$.

Exercise 5. From the expressions (4.63) and (4.64) for u and $\boldsymbol{V}$ and from eq. (4.36) we have

$$\begin{aligned} \frac{\partial u}{\partial t} + c_o^2 div\mathbf{V} &= \frac{1}{4\pi\epsilon_o}\int_0^{c_o t}\frac{dr}{r}\int_{\Sigma_r}\rho_t(\mathbf{y}, t-\frac{r}{c_o})\,dS_y + c_o^2\frac{\mu_o}{4\pi}\int_0^{c_o t}\frac{dr}{r}\int_{\Sigma_r} div\mathbf{J}_r(\mathbf{y}, t-\frac{r}{c_o})\,dS_y \\ &\equiv \frac{1}{4\pi\epsilon_o}\int_0^{c_o t}\frac{dr}{r}\int_{\Sigma_r}(\rho_t - div\mathbf{J}_r)dS_y \;=\; 0 \end{aligned}$$

since $c_o^2\mu_o = 1/\epsilon_o$.

Exercise 6. For a transmission line made up of two parallel straigth wire conductors with circular cross section of radius r_o , distance d and length l, the inductance per unit length L' is defined as

$$L' := 2(L_{11} - L_{12})/l$$

where $L_{11} = L_{22}$ and L_{12} are the self-inductance and mutual inductance of the two wires, respectively [43]. Show that for $l \gg d \gg r_o$, L' is given by

$$L' \cong \frac{\mu_o}{\pi} \, log \, \frac{d}{r_o}$$

Hint: see Exercise 9 of Chapter 3.

Exercise 7. Verify that (4.76) is solution to eq. (4.75) for $w(x)$ any regular function of x.

Exercise 8. We have

$$I = e^{-\beta t} w(x - at) \equiv \; e^{-\beta x/a} e^{\beta(x-at)/a} w(x - at) \equiv e^{-\beta x/a} W(x - \, at)$$

where

$$W(x) := e^{\beta x/a} w(x)$$

is an arbitrary function, since so is w.

Exercise 9. Verify eq. (4.78).

Exercise 10. If $\boldsymbol{u}$ is a weak solution in $C^1(R_T)$ and $\boldsymbol{w} \in C^o(R_T)$, taking any $\boldsymbol{v}(x_o,\boldsymbol{x}) \in C^1_o(R_T)$ with $\boldsymbol{v}(0,\boldsymbol{x}) = 0$ in eq. (4.92) and integrating by parts backwards for $x_o > 0$ we find that the equation

$$((L^*\boldsymbol{v}, \boldsymbol{u})) - \, ((\boldsymbol{v}, \boldsymbol{w})) = ((\boldsymbol{v}, L\boldsymbol{u} - \boldsymbol{w})) = 0$$

must be satisfied with any such $\boldsymbol{v}$. A well known result from Calculus then implies that $L\boldsymbol{u}$=$\boldsymbol{w}$, and so $\boldsymbol{u}$ is a classical solution of the differential system for $x_o > 0$.

Next, taking $\boldsymbol{v} \in C^1_o(R_T)$ with $\boldsymbol{v}(0,\boldsymbol{x}) \neq 0$ we have, denoting $(\boldsymbol{v},\boldsymbol{u})$ the scalar product in $L^2(\mathbb{R}^3)$,

$$((L^*\boldsymbol{v}, \boldsymbol{u})) - \, ((\boldsymbol{v}, \boldsymbol{w})) = ((\boldsymbol{v}, L\boldsymbol{u} - \boldsymbol{w})) + (\boldsymbol{v}, \boldsymbol{u})_{x_o=0} \; = (\boldsymbol{v}, \boldsymbol{u})_{x_o=0} \; = 0$$

and since $\boldsymbol{v}$ is arbitrary the homogeneous initial condition $\boldsymbol{u}(0,\boldsymbol{x}) = 0$ is also satisfied.

Exercise 11. Derive eq. (4.111).

Exercise 12. Derive eqs. (4.117).

Exercise 13. (i) The eikonal function $\mathcal{I} = \boldsymbol{k}\cdot\boldsymbol{x}$ satisfies the eikonal equation (4.120) for every unit vector $\boldsymbol{k}$. The corresponding characteristics are families of planes

$$\Phi = \mathbf{k} \cdot \mathbf{x} \pm ct = \text{ constant}$$

and the wavefronts are planes in $\mathbb{R}^3$ with normal $\boldsymbol{n}=\boldsymbol{k}$ for any fixed t moving with normal speed $\pm c$:therefore the surfaces of constant phase of the plane waves (4.10) are wavefronts. The bicharacteristic rays $\boldsymbol{x}-\boldsymbol{x}_o = \pm c\boldsymbol{k}t$ are straight lines lying on the characteristic surfaces and, viewed in $\mathbb{R}^3$, are trajectories orthogonal to the wavefronts.

(ii) In particular for $\boldsymbol{k}= (1,0,0)$ and $\mathcal{I} = x_1$ the wavefronts $x_1 = \pm ct$ =constant are planes orthogonal to the x_1-axis in $\mathbb{R}^3$ for any fixed t, and the bicharacteristic rays viewed in $\mathbb{R}^3$ are straight lines parallel to the x_1-axis.

(iii) The eikonal function $\mathcal{I} = |\boldsymbol{x}-\boldsymbol{x}_o|$ satisfies the eikonal equation (4.120) for every fixed $\boldsymbol{x}_o \in \mathbb{R}^3$. The corresponding characteristics are families of cones

$$\Phi = |\mathbf{x} - \mathbf{x}_o| \pm ct = R = \text{constant}$$

and the wavefronts are moving spheres with radii $R\mp ct$. The bicharacteristic rays are the generatrices of the characteristic cones in $\mathbb{R}^4$, and viewed in $\mathbb{R}^3$ they are orthogonal trajectories of the wavefronts, i.e. bundles of straight lines centered at $\boldsymbol{x}_o$.

Exercise 14. Verify that the general solution of (4.138) holds.

Hint: Dividing the second equation (4.138) by h_i and adding the first yields

$$\frac{\partial}{\partial\tau}(h_i u_i \pm v_i) \pm \frac{1}{h_i}\frac{\partial}{\partial z}(h_i u_i \pm v_i) = 0$$

Hence the functions $U_i := h_i u_i + v_i$ are constant along the lines $C_+ : \tau - h_i z = const.$, whereas the functions $V_i := h_i u_i - v_i$ are constant along the lines $C_- : \tau + h_i z = const.$ More generally, in any convex set of $\mathbb{R}^2$ we have

$$U_i = \Phi(\tau - h_i z) \quad , \quad V_i = \Psi(\tau + h_i z)$$

where Φ, Ψ are arbitrary C^1 functions [2]. In the theory of partial differential equations C_+, C_- are the characteristic curves of the hyperbolic system (4.138).

From all the above considerations it follows that the solutions u_i, v_i are given by the superpositions of Riemann invariants

$$u_i(z,\tau) = \frac{1}{h_i}(U_i(\tau - h_i z) + V_i(\tau + h_i z)) \,, \quad v_i(z,\tau) = U_i(\tau - h_i z) - V_i(\tau + h_i z)$$

as asserted.

Exercise 15. (i) Show that the boundary conditions (4.141), (4.143) can be written in terms of E, H in the form

$$\text{(E3)} \qquad E(0,t) + \alpha H(0,t) = I(t) \;, \;\; E(a,t) - \beta H(a,t) = 0$$

where $\alpha = \sqrt{\mu_o/\varepsilon_1}\,$, $\beta = \sqrt{\mu/\varepsilon_3}\,$, $I(t) = 2\sqrt{\mu_o/\varepsilon_1}\,\mathcal{I}(\tau/c_2)$ (α and β are wave impedances, see §4.2.1).

(ii) The energy corresponding to a solution (E,H) of the BVP (4.145) in D_a is defined by

$$\mathcal{E}(t) = \frac{1}{2}\int_0^a (\varepsilon_2 E^2(x,t) + \mu_o\, H^2(x,t))\, dx$$

If $E, H \,\epsilon\, C^1(\overline{D}_a)$, show that $\mathcal{E}'(t) < 0$ for $I(t) \equiv 0$ (the boundary conditions (E3) are dissipative as t increases). Thus for $\alpha > 0, \beta > 0$ the energy may diverge as $t \to -\infty$. If α and β were negative, or if the signs in (E3) are reversed, this would take place as $t \to +\infty$.

[Hint: $\mathcal{E}'(t) \le -\beta H^2(a,t) - \alpha H^2(0,t) < 0$.]

Exercise 16. Let V_o satisfy the homogeneous difference equation associated to eq. (4.149)

$$\text{(E4)} \quad V_o(\tau + 2a) + r_1 r_2 V_o(\tau) = 0$$

and put $P(\tau) := |r_1 r_2|^{-\tau/2a} V_o(\tau)$. Then the function $P(\tau)$ satisfies

$$P(\tau + 2a) \,+\, sgn(r_1 r_2)\, P(\tau) = 0$$

and is otherwise arbitrary. Therefore the solution of (E4) is

$$V_o(\tau) = |r_1 r_2|^{\tau/2a} P(\tau)$$

where $P(\tau)$ is any periodic bounded function, with period $2a$ (if $r_1r_2 < 0$) or antiperiod $2a$ (if $r_1r_2 > 0$). It follows that

(i) If $|r_1r_2| = 1$ (E4) has (infinitely many) bounded solutions $V_o(\tau)$, and so a bounded solution to eq. (4.149) is not unique.

(ii) If $|r_1r_2| \neq 1$, all solutions $V_o(\tau)$ of (E4) are unbounded for $\tau \in \mathbb{R}$, hence (4.149) has at most one bounded solution.

In the case of eq. (4.149) we have $|r_1r_2| < 1$ and the series (4.150) converges uniformly, since

$$\left|\sum_{n=0}^{\infty} (-r_1r_2)^n F(\tau - 2a - 2na)\right| \leq \sum_{n=0}^{\infty} |-r_1r_2|^n \sup_{\tau\in\mathbb{R}} |F(\tau)|$$
$$= \frac{1}{1-|r_1r_2|} \frac{2|r_2|}{1+h_1} \sup_{t\in\mathbb{R}} |\mathcal{I}(t)| \qquad (\forall \tau \in \mathbb{R})$$

It follows that $V(\tau)$, the sum of the series (4.150), is bounded and satisfies (4.149):

$$V(\tau+2a) + r_1r_2V(\tau) = \sum_{n=0}^{\infty} (-r_1r_2)^n (F(\tau-2na) - r_1r_2F(\tau-2a-2na))$$
$$= F(\tau) + \sum_{n=1}^{\infty} (-r_1r_2)^n (F(\tau-2na) - F(\tau-2na)) = F(\tau)$$

Exercise 17. Show that

(i) The solution of BVP depends continuously on $\mathcal{I}(\tau)$.

(ii) If $\mathcal{I}(\tau)$ has compact support in $[\tau_o, \tau_1]$, the support of $\mathcal{R}(\tau)$ is in $[\tau_o, \infty)$ and $\mathcal{R}(\tau) \to 0$ as $\tau \to +\infty$. (The solution decays as $\tau \to \infty$. This corresponds to the physical fact that the wave amplitude is reduced at each reflection.)

(iii) For $\mathcal{I}(\tau) \equiv 0$ the homogeneous BVP has the unbounded solutions

$$u(z,\tau) = r_1|r_1r_2|^{(\tau-z)/2a} P(\tau - z) + |r_1r_2|^{(\tau+z)/2a} P(\tau + z)$$

$$v(z,\tau) = r_1|r_1r_2|^{(\tau-z)/2a} P(\tau - z) - |r_1r_2|^{(\tau+z)/2a} P(\tau + z)$$

where P is (arbitrary) periodic with $P(\tau + 2a) + sgn(r_1r_2)\, P(\tau) = 0$. (Cf. Exercise 15).

Exercise 18. Derive eq. (4.162). Find the expressions of the reflected and transmitted waves $\mathcal{R}(\tau)$, $\mathcal{T}(\tau)$ for two layers.

Exercise 19. The homogeneous equation

$$V(\tau) - A_1 V(\tau + b_1) - A_2 V(\tau + b_2) = 0 \quad (\tau \in \mathbb{R})$$

with $N = 2$, $b_1 = \varepsilon$, $0 < \varepsilon \ll 1$, $b_2 = ln(3) + O(\varepsilon)$, $A_1 = A_2 = 1/4$, has the unbounded solution $V_o = exp(\tau)$. Here assumption (4.183) is satisfied, with $\beta = 0$.

Exercise 20. The homogeneous equation with $N = 2$

$$V(\tau) + \frac{1}{\sqrt{2}} V(\tau + \frac{3\pi}{4}) + \frac{1}{\sqrt{2}} V(\tau + \frac{5\pi}{4}) = 0 \quad (\tau \in \mathbb{R})$$

has the bounded solutions $V_o = sin\,\tau$, $V_o = cos\,\tau$. Here assumption (4.183) is not satisfied.

Chapter 5

Electrodynamics Of Moving Bodies

The Maxwell equations for a homogeneous medium considered in previous chapters

$$\frac{\partial \boldsymbol{B}}{\partial t} = -curl\, \boldsymbol{E} \tag{5.1}$$

$$\frac{\partial \boldsymbol{D}}{\partial t} = curl\, \boldsymbol{H} - \boldsymbol{J} \tag{5.2}$$

$$div\, \boldsymbol{D} = \rho\,,\ div\, \boldsymbol{B} = 0 \tag{5.3}$$

$$\boldsymbol{D} = \epsilon \boldsymbol{E}\ ,\quad \boldsymbol{B} = \mu \boldsymbol{H} \tag{5.4}$$

are referred to a privileged observer $O(x_1, x_2, x_3, t)$ which in practice must be at rest with respect to the laboratory. Suppose $O'(x'_1, x'_2, x'_3, t')$ is another inertial observer moving with constant velocity $\boldsymbol{v}$ with respect to O. In the transition from O to O' the Galilei transformations should hold, and the Maxwell equations should be invariant with respect to the Galilei group, as it happens for the Newton equation of mechanics. If $\boldsymbol{v}= v\mathbf{c}_1$ the Galilei transformations read

$$t' = t,\ x'_1 = x_1 - vt,\ x'_2 = x_2,\ x'_3 = x_3 \tag{5.5}$$

It is easy to verify, however, that this is not true: the Maxwell equations are not Galilei-invariant, even in the tensorial sense of shape invariance. In fact, due to the presence of the displacement current, the Maxwell equations are invariant in shape with respect to a different group of linear transformations, discovered by H. A. Lorentz at the end of the 19th century. For $\boldsymbol{v}= v\mathbf{c}_1$ they have the form

$$t' = \alpha t - \frac{\alpha\beta}{c} x_1 \ , \ \ x_1' = \alpha x_1 - \alpha v t \ , \ x_2' = x_2 \ , \ x_3' = x_3 \tag{5.6}$$

where

$$\alpha := \frac{1}{\sqrt{1-\beta^2}} \ , \quad \beta := \frac{v}{c} \tag{5.7}$$

and $c = (\epsilon\mu)^{-\frac{1}{2}}$ is the velocity of light. The Lorentz transformations (5.6), (5.7) involve also a change of time, and reduce to the Galilei transformations (5.5) only in the limit $c \to \infty$.

Thus either the Galilei transformations are discarded, or a privileged observer for the electromagnetic phenomena has to be singled out. Historically, the second alternative was chosen and a solution was attempted by introducing the ad hoc concept of (luminiferous) ether, a hypothetical medium pervading all space and conceived both as the seat of electromagnetic phenomena and as the support of electromagnetic waves, whose propagation in vacuo was then regarded as a physical impossibility. In this context, the privileged observer was envisaged as being fixed with respect to the ether. However, the overwhelming experimental evidence which subsequently accumulated led A. Einstein to publish in 1905 a celebrated paper [1] in which the ether hypothesis was rejected and the Maxwell equations were extended to moving bodies by framing the electromagnetic laws in space-time. In this framework, Einstein also discarded the Galilei trasformations and reformulated the laws of mechanics so as to render them compatible with the Lorentz group [43]. The outcome of this process was the well-known model of the physical world called "restricted Relativity theory", which furnishes the natural context of Maxwell's electromagnetism.

[1] A.Einstein: Zur Elektrodynamik bewegter Körper, Annalen der Physik 17 (1905)

5.1 Lorentz invariance of the Maxwell equations

Setting

$$x_o := ict$$

and introducing the four-dimensional complex Minkowski space-time, or chronotope $(x_o, ..., x_3)$, endowed of the pseudo-euclidean metric

$$ds^2 := \sum_{k=0}^{3} dx_k^2 = dx_1^2 + dx_2^2 + dx_3^2 - c^2dt^2 \tag{5.8}$$

the special Lorentz transformation (5.6) takes the complex form

$$x'_o = \alpha x_o - i\alpha\beta x_1 \ , \ \ x'_1 = i\alpha\beta x_o + \alpha x_1 \, , \ \ x'_2 = x_2 \ , \ x'_3 = x_3 \tag{5.9}$$

It is easy to verify (Exercise 1) that while the metric (5.8) is invariant with respect to Lorentz transformations, the spatial distances and the time durations change [2]. If we introduce the complex angle

$$\theta = arctan\,(i\beta) \equiv i\psi$$

we have

$$sin\theta = \frac{i\beta}{\sqrt{1-\beta^2}} = i\alpha\beta \equiv isinh\,\psi \quad , \qquad cos\,\theta = \frac{1}{\sqrt{1-\beta^2}} = \alpha \equiv cosh\,\psi$$

and eq. (5.9) takes the form of the imaginary rotation of the chronotope

$$x'_o = x_o cos\,\theta - x_1 sin\theta \ , \ \ x'_1 = x_o sin\theta + x_1 cos\,\theta \, , \ \ x'_2 = x_2 \ , \ x'_3 = x_3$$

or equivalently of the hyperbolic rotation in the $4D$ euclidean space (x_1, x_2, x_3, ct) :

$$ct' = ct\,cosh\,\psi - x_1 sinh\,\psi \ , \ \ x'_1 = -ctsinh\,\psi + x_1 cosh\,\psi \, , \ \ x'_2 = x_2 \ , x'_3 = x_3$$

Since $(1-\beta^2)\alpha^2 = 1$, the 4×4 Jacobian matrix

$$\left[\frac{\partial \mathbb{X}'}{\partial \mathbb{X}}\right] = \left[\frac{\partial x'_k}{\partial x_l}\right]_{k,l=0,...3} \quad ; \qquad \mathbb{X} = (x_o, ..., x_3)\, , \ \mathbb{X}' = (x'_o, ..., x'_3)$$

[2] hence the well-known relativistic phenomena of "contraction of lengths" and "dilation of times" [43]

for the Lorentz transformation (5.9) takes the form

$$\left[\frac{\partial x'_k}{\partial x_l}\right] = \begin{pmatrix} \alpha & -i\alpha\beta & 0 & 0 \\ i\alpha\beta & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.10}$$

so that its determinant equals 1 and the inverse and transposed matrices coincide. Therefore the inverse Lorentz transformation reads

$$x_o = \alpha x'_o + i\alpha\beta x'_1 \ , \ \ x_1 = -i\alpha\beta x'_o + \alpha x'_1 \, , \ x_2 = x'_2 \ , \ x_3 = x'_3$$

We denote by $\|$,$\perp$ the components of a vector parallel and orthogonal to $\boldsymbol{v}$, respectively: for example

$$\boldsymbol{J}_\| := \boldsymbol{J}\cdot\boldsymbol{v}\ \frac{\boldsymbol{v}}{|\boldsymbol{v}|^2}\, , \ \ \boldsymbol{J}_\perp := -(\boldsymbol{J}\wedge\boldsymbol{v})\wedge\frac{\boldsymbol{v}}{|\boldsymbol{v}|^2} \equiv \boldsymbol{J}-\boldsymbol{J}_\|$$

so that $(\boldsymbol{J}\wedge\boldsymbol{v})_\| \equiv \boldsymbol{0}$, $(\boldsymbol{J}\wedge\boldsymbol{v})_\perp \equiv \boldsymbol{J}\wedge\boldsymbol{v}$. Then the general Lorentz transformation for an arbitrary $\boldsymbol{v}$ is

$$t' = \alpha(t - \frac{\boldsymbol{v}\cdot\boldsymbol{x}}{c^2})\, , \quad \boldsymbol{x}' = \alpha(\boldsymbol{x}-\boldsymbol{v}t)_\| + \boldsymbol{x}_\perp \equiv \alpha(\boldsymbol{x}-\boldsymbol{v}t)_\| + (\boldsymbol{x}-vt)_\perp \tag{5.11}$$

and corresponds to the passage from the observer $O(x_1, x_2, x_3, t)$ to an observer $O'(x'_1, x'_2, x'_3, t')$ moving with constant velocity $\boldsymbol{v}$ with respect to O. Let us define new field quantities $\boldsymbol{E}'$,$\boldsymbol{B}'$, ρ' and $\boldsymbol{J}'$ as

(5.12)

$$\boldsymbol{E}' = (1-\alpha)\boldsymbol{E}_\| + \alpha(\boldsymbol{E}+\boldsymbol{v}\wedge\boldsymbol{B}) \equiv (\boldsymbol{E}+\boldsymbol{v}\wedge\boldsymbol{B}) + \alpha(\boldsymbol{E}+\boldsymbol{v}\wedge\boldsymbol{B})_\perp$$

$$\boldsymbol{B}' = (1-\alpha)\boldsymbol{B}_\| + \alpha(\boldsymbol{B} - \frac{1}{c^2}\boldsymbol{v}\wedge\boldsymbol{E}) \equiv (\boldsymbol{B} - \frac{1}{c^2}\boldsymbol{v}\wedge\boldsymbol{E})_\| + \alpha(\boldsymbol{B} - \frac{1}{c^2}\boldsymbol{v}\wedge\boldsymbol{E})_\perp$$

$$\rho' = \alpha(\rho - \frac{\boldsymbol{v}\cdot\boldsymbol{J}}{c^2})\, , \quad \boldsymbol{J}' = \alpha(\boldsymbol{J}-\rho\boldsymbol{v})_\| + \boldsymbol{J}_\perp \equiv \alpha(\boldsymbol{J}-\rho\boldsymbol{v})_\| + (\boldsymbol{J}-\rho\boldsymbol{v})_\perp$$

(the interpretation of these equations will be given later on), and let

$$\boldsymbol{D}' = \epsilon\boldsymbol{E}'\ ,\ \ \boldsymbol{H}' = \mu^{-1}\boldsymbol{B}'$$

Since $c = (\epsilon\mu)^{-\frac{1}{2}}$, the explicit expressions for $\boldsymbol{D}'$ and $\boldsymbol{H}'$ are

$$\boldsymbol{D}' = (\boldsymbol{D}+\frac{1}{c^2}\boldsymbol{v}\wedge\boldsymbol{H})_\| + \alpha(\boldsymbol{D}+\frac{1}{c^2}\boldsymbol{v}\wedge\boldsymbol{H})_\perp\, , \ \ \boldsymbol{H}' = (\boldsymbol{H}-\boldsymbol{v}\wedge\boldsymbol{D})_\| + \alpha(\boldsymbol{H}-\ \boldsymbol{v}\wedge\boldsymbol{D})_\perp$$

If (without loss of generality) we assume that $\boldsymbol{v}= v\mathbf{c}_1$, eqs.(5.12) reduce to

(5.13)
$$\begin{aligned}
&E_1' = E_1,\ E_2' = \alpha E_2 - \alpha\beta c B_3,\ E_3' = \alpha E_3 + \alpha\beta c B_2\\
&B_1' = B_1,\ B_2' = \alpha B_2 + \alpha\beta c^{-1} E_3\,,\ B_3' = \alpha B_3 - \alpha\beta c^{-1} E_2\\
&\rho' = \alpha\rho - \alpha\beta c^{-1} J_1\ ,\quad J_1' = \alpha J_1 - \alpha\beta c\rho,\quad J_2' = J_2\,,\ J_3' = J_3
\end{aligned}$$

Proposition 5.1.1 *(invariance of the Maxwell equations under Lorentz transformations). Under the transformations (5.11), the Maxwell equations (5.1)–(5.3) maintain the same form*

$$\begin{aligned}
&\frac{\partial \boldsymbol{B}'}{\partial t} = -curl\ \boldsymbol{E}'\\
&\frac{\partial \boldsymbol{D}'}{\partial t} = curl\ \boldsymbol{H}' - \boldsymbol{J}'\\
&div\ \boldsymbol{D}' = \rho'\ div \boldsymbol{B}' = 0
\end{aligned}$$

provided $\boldsymbol{E}', \boldsymbol{B}', \boldsymbol{D}' = \epsilon \boldsymbol{E}', \boldsymbol{H}' = \mu^{-1}\boldsymbol{B}', \rho'$ *and* $\boldsymbol{J}'$ *are defined according to (5.12) or (5.13).*

Proof. See Exercise 2.

This result implies that the Maxwell equations are not Galilei invariant (Exercise 2).

Eqs. (5.12) and (5.13) show that in the passage from a fixed to a moving observer the 3D field vectors $\boldsymbol{E}$ and $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{H}$ merge to form two hexa-vectors $(\boldsymbol{E},\boldsymbol{B})$ and $(\boldsymbol{D},\boldsymbol{H})$ in the chronotope, that is, two quantities consisting of two 3D vectors each, corrisponding to two 4-D emisymmetric double tensors. Similarly, the charge density ρ and the current $\boldsymbol{J}$ combine to form a tetra-vector with four scalar components. This mirrors the fact that an electrostatic field in the laboratory reference frame becomes an electromagnetic field if seen from a moving observer, and that a moving charge generates a current flow called convection current. The first eq. (5.12) shows that $\boldsymbol{E}$ and $\boldsymbol{B}$ combine in the form $\boldsymbol{E}+\boldsymbol{v}\wedge\boldsymbol{B}$ that, multiplied by the charge, furnishes the expression of the Lorentz force

$$\mathbf{F} = q(\boldsymbol{E} + \boldsymbol{v} \wedge \boldsymbol{B})$$

exerted by an electromagnetic field $\boldsymbol{E}$, $\boldsymbol{B}$ on a pointwise particle of charge q moving with velocity $\boldsymbol{v}$. This force is the sum of the force $q\boldsymbol{E}$ due to the electric field (see eq. (1.1)) and of the deviating (power-free) force

$$q\boldsymbol{v} \wedge \boldsymbol{B}$$

due to the magnetic field. The latter agrees with the Ampere rule of Chapter 1 if the current density due to the particle motion is defined as

$$\boldsymbol{J} := \rho\boldsymbol{v}$$

and if the corresponding charge density is defined by

$$\rho = q\delta(\boldsymbol{x} - \boldsymbol{x}_q)$$

where $\delta(\boldsymbol{x}-\boldsymbol{x}_q)$ is the Dirac distribution (measure) in $\mathbb{R}^3$ centered at the particle location $\boldsymbol{x}_q$ (Exercise 3).

We are thus led to the following conclusions:

(i) The natural setting of the electromagnetic field is the chronotope, endowed of the metric (5.8).

(ii) The passage from an "inertial" observer to another "inertial" observer in uniform translational motion with respect to the first is described by the Lorentz trasformations of special Relativity theory. These transformations imply that time and length are relative concepts and depend on the observer.

(iii) All field quantities must be represented in tensor form by means of scalars, tetra-vectors and emisymmetric tensors (hexa-vectors), so as to be invariant [3] with respect to Lorentz transformations.

In other words, from an algebraic point of view the field quantities are matrix elements, with respect to a cartesian basis, of (multi)linear applications of the chronotope viewed as a 4-D linear space, and the Lorentz transformations correspond to change of bases in the chronotope. Consider a tetra-vector, say $\mathbb{J} = \mathbb{J}(\mathbb{X})$ for a given system of cartesian coordinates $\mathbb{X} = (x_o, ..., x_3)$. Denoting by $\mathbb{J}_k$, $\mathbb{J}'_l$ the contravariant components of $\mathbb{J}$ in the basis $\left\{\frac{\partial \mathbb{X}}{\partial x_k}\right\}$ and $\left\{\frac{\partial \mathbb{X}}{\partial x'_l}\right\}$, respectively, we have

$$\mathbb{J} = \sum_{k=0}^{3} \mathbb{J}_k \frac{\partial \mathbb{X}}{\partial x_k} = \sum_{k,l=0}^{3} \mathbb{J}_k \frac{\partial x'_l}{\partial x_k} \frac{\partial \mathbb{X}}{\partial x'_l} = \sum_{l=0}^{3} \mathbb{J}'_l \frac{\partial \mathbb{X}}{\partial x'_l}$$

[3] more precisely, covariant or contravariant, see what follows

and so the new contravariant components $\mathbb{J}'_l$ in the new system of coordinates $\mathbb{X}' = (x'_o, ..., x'_3)$ are defined in terms of the old contravariant components $\mathbb{J}_k$ by the linear relations

$$\mathbb{J}'_l = \sum_{m=0}^{3} \frac{\partial x'_l}{\partial x_k} \mathbb{J}_k \qquad (l = 0, ..., 3) \tag{5.14}$$

Similarly, if $\mathbb{F}$ is a double tensor we have

$$\mathbb{F} = \sum_{j,k=0}^{3} \mathbb{F}_{j,k} \frac{\partial \mathbb{X}}{\partial x_j} \otimes \frac{\partial \mathbb{X}}{\partial x_k} = \sum_{l,m=0}^{3} \mathbb{F}'_{l,m} \frac{\partial \mathbb{X}'}{\partial x_l} \otimes \frac{\partial \mathbb{X}'}{\partial x_m}$$

where

$$\mathbb{F}'_{l,m} = \sum_{j,k=0}^{3} \frac{\partial x'_l}{\partial x_j} \frac{\partial x'_m}{\partial x_k} \mathbb{F}_{lm} \qquad (l, m = 0, ..., 3) \tag{5.15}$$

These relations, eq. (5.14) and (5.15), are the standard formulae for the change of coordinates of contravariant components of vectors and tensors in a 4D (complex pseudo-) euclidean space. The corresponding formulae for covariant components are obtained by exchanging x_k and x'_l. The particular change of coordinates described by the Lorentz transformations (5.11) (i.e. (5.9) for $\boldsymbol{v} = v\mathbf{c}_1$) corresponds physically to the passage from the observer O to the observer O' translating uniformly with velocity $\boldsymbol{v}$ with respect to O.

5.2 The Maxwell equations in four dimensions

The electromagnetic field vectors are given in the chronotope by two hexa-vectors, or emisymmetric double tensors $\mathbb{F}$ and $\mathbb{G}$, whose (contravariant) components in the system of coordinates $\mathbb{X} = (x_o, ..., x_3)$ are represented by the complex matrix elements [4]

$$\mathbb{F}_{jk} := \begin{pmatrix} 0 & iE_1 & iE_2 & iE_3 \\ -iE_1 & 0 & cB_3 & -cB_2 \\ -iE_2 & -cB_3 & 0 & cB_1 \\ -iE_3 & cB_2 & -cB_1 & 0 \end{pmatrix}$$

[4] a representation in terms of real matrix elements is also possible

$$\mathbb{G}_{jk} := \begin{pmatrix} 0 & icD_1 & icD_2 & icD_3 \\ -icD_1 & 0 & H_3 & -H_2 \\ -icD_2 & -H_3 & 0 & H_1 \\ -icD_3 & H_2 & -H_1 & 0 \end{pmatrix}$$

The constitutive relations (5.4) take the form

$$\mathbb{G} = \sqrt{\frac{\epsilon}{\mu}}\,\mathbb{F} \quad \Rightarrow \quad \mathbb{G}_{jk} = \sqrt{\frac{\epsilon}{\mu}}\,\mathbb{F}_{jk} \qquad (j,k=0,...,3)$$

The matrix elements of $\mathbb{F}$ and $\mathbb{G}$ in a new coordinate system $O'(x'_o, x'_1, x'_2, x'_3)$ are given by

(5.16)

$$\mathbb{F}'_{jk} = \sum_{l,m=0}^{3} \frac{\partial x'_j}{\partial x_l}\frac{\partial x'_k}{\partial x_m}\,\mathbb{F}_{lm} \quad , \quad \mathbb{G}'_{jk} = \sum_{l,m=0}^{3} \frac{\partial x'_j}{\partial x_l}\frac{\partial x'_k}{\partial x_m}\,\mathbb{G}_{lm} \quad (j,k=0,...,3)$$

(see eqs. (5.15)). The “current density” $\mathbb{J}$ forms a tetra-vector in the chronotope whose contravariant components in the system of coordinates $\mathbb{X} = (x_o, ..., x_3)$ are given by

(5.17) $$\mathbb{J} := \ (ic\rho,\ \mathrm{J}) = (ic\rho\ ,\ J_1, J_2, J_3) \quad (j=0,...,3)$$

and the “tetra-potential” $\mathbb{U}$ is defined by

(5.18) $$\mathbb{U} := (i\frac{u}{c},\ \boldsymbol{V}) = (i\frac{u}{c},\ V_1, V_2, V_3) \quad (j=0,...,3)$$

For a change of coordinates the contravariant components of $\mathbb{J}$ and $\mathbb{U}$ in the new coordinate system $O'(x'_o, x'_1, x'_2, x'_3)$ are

(5.19) $$\mathbb{J}'_j = \sum_{m=0}^{3} \frac{\partial x'_j}{\partial x_m}\,\mathbb{J}_m\ , \quad \mathbb{U}'_j = \sum_{m=0}^{3} \frac{\partial x'_j}{\partial x_m}\,\mathbb{U}_m \qquad (j=0,...,3)$$

(see eq. (5.14)). In the case of the special Lorentz transformation (5.9) the Jacobian matrix $\left[\frac{\partial x'_j}{\partial x_l}\right]$ is given by (5.10), eqs. (5.16), (5.19) reduce to (5.13), and for the tetra-potential we find

(5.20)

$$u' = \alpha(u - \boldsymbol{v}\cdot\boldsymbol{V})\,, \quad \boldsymbol{V}' = \alpha(\boldsymbol{V} - \frac{u}{c^2}\boldsymbol{v})_{\parallel} + \boldsymbol{V}_{\perp} \equiv \alpha(\boldsymbol{V} - \frac{u}{c^2}\boldsymbol{v})_{\parallel} + (\boldsymbol{V} - \frac{u}{c^2}\boldsymbol{v})_{\perp}$$

(Exercise 4). Note that the scalar and vector potentials combine in these expressions to form the quantity $u - \boldsymbol{v} \cdot \boldsymbol{V}$ which multiplied by ρ and taking $\boldsymbol{J} := \rho \boldsymbol{v}$ defines the so-called Schwartzschild invariant

$$\mathbb{J} \cdot \mathbb{U} = -\rho u + \boldsymbol{J} \cdot \boldsymbol{V}$$

Note also that the charge density ρ is the first component of $\mathbb{J}$ and hence depends on the basis (system of coordinates) chosen. Thus ρ is not invariant with respect to a Lorentz transformation. However, the charge of a particle (in particular the charge e of the electron) is an absolute invariant [5], i.e. is invariant with respect to any change of coordinates in the chronotope (we omit the proof).

Let us denote by * the switch operator which exchanges the real and imaginary parts of any double tensor: for example, $^*\mathbb{F}$ has matrix elements given by

$$^*\mathbb{F}_{jk} = \begin{pmatrix} 0 & cB_1 & cB_2 & cB_3 \\ -cB_1 & 0 & iE_3 & -iE_2 \\ -cB_2 & -iE_3 & 0 & iE_1 \\ -cB_3 & iE_2 & -iE_1 & 0 \end{pmatrix}$$

We also define the ad hoc differential operators Rot, $\mathbf{Div}$,Div,$Grad$,$\Box$ in a cartesian system of coordinates $\mathbb{X} = (x_o, ..., x_3)$ as follows [43]. The operator Rot is defined on tetra-vectors and produces a tensor according to the componentwise definition in the chosen basis

$$(Rot\,\mathbb{U})_{jk} := \frac{\partial \mathbb{U}_k}{\partial x_j} - \frac{\partial \mathbb{U}_j}{\partial x_k} \qquad (j, k = 0, ..., 3)$$

The operator $\mathbf{Div}$ is defined on tetra-vectors and produces a tetra-vector according to

$$(\mathbf{Div}\,\mathbb{G})_j := \sum_{k=0}^{3} \frac{\partial \mathbb{G}_{jk}}{\partial x_k} \qquad (j = 0, ..., 3)$$

The operator Div is the usual four-dimensional scalar divergence, defined on tetra-vectors by

$$Div\,\mathbb{J} := \sum_{k=0}^{3} \frac{\partial \mathbb{J}_k}{\partial x_k} \qquad (j = 0, ..., 3)$$

[5] incidentally, in the restricted Relativity theory the mass depends on the speed of the particle and hence is not invariant [43]

The operator *Grad* is the usual four-dimensional gradient defined on scalars f :

$$Gradf := (\frac{\partial f}{\partial x_o}, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3})$$

Last, the operator $\Box$ is the usual four-dimensional Laplace operator in the coordinates (x_o, x_1, x_2, x_3), or the wave operator in the coordinates (t, x_1, x_2, x_3) :

$$\Box := \sum_{k=0}^{3} \frac{\partial^2}{\partial x_k^2} = \Delta_3 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$$

These operators satisfy the following identities:

(5.21) $$Div\, \mathbf{Div} \equiv 0 \quad , \quad \mathbf{Div}\, Rot \equiv Grad\, Div - \Box \quad , \quad \mathbf{Div}\, {}^*Rot \equiv 0$$

(Exercise 5). The Maxwell equations in four-dimensional form read then

(5.22) $$\mathbf{Div}\, \mathbb{G} = \mathbb{J} \quad , \quad \mathbf{Div}\, {}^*\mathbb{F} = 0$$

and the constitutive equations for a homogenous non-magnetic medium are

(5.23) $$\mathbb{G} = \sqrt{\frac{\epsilon}{\mu}}\ \mathbb{F}$$

These equations combined with the identities (5.21) imply that:

(i) $\mathbb{J}$ is solenoidal:

(5.24) $$Div\, \mathbb{J} = 0$$

(ii) the general solution of the second eq. (5.22) has the form

(5.25) $$\mathbb{F} = c\, Rot\, \mathbb{U}$$

where $\mathbb{U}$ is interpreted as the tetra-potential

(iii) in a homogeneous non-magnetic medium $\mathbb{G}$ is related to $\mathbb{U}$ by

(5.26) $$\mathbb{G} = \mu^{-1}\, Rot\, \mathbb{U}$$

Moreover, the tetra-potential $\mathbb{U}$ satisfies

(5.27) $$Div\, \mathbb{U} = 0 \quad , \quad \Box\mathbb{U} = -\mu_o\, \mathbb{J}$$

Indeed, the following correspondence table is easy to verify:

$$\mathbf{Div}\,\mathbb{G} = \mathbb{J} \quad \Leftrightarrow \quad div\boldsymbol{D} = \rho \ , \ \frac{\partial \boldsymbol{D}}{\partial t} = curl\,\boldsymbol{H} - \boldsymbol{J}$$

$$\mathbf{Div}\,{}^*\mathbb{F} = 0 \quad \Leftrightarrow \quad div\boldsymbol{B} = 0 \ , \ \frac{\partial \boldsymbol{B}}{\partial t} = -curl\,\boldsymbol{E}$$

$$\mathbb{F} = c\,Rot\,\mathbb{U} \quad \Leftrightarrow \quad \boldsymbol{B} = curl\ \boldsymbol{V} \quad , \qquad \boldsymbol{E} = -grad\,u - \frac{\partial \boldsymbol{V}}{\partial t}$$

$$\Box\,\mathbb{U} = -\mu\,\mathbb{J} \quad \Leftrightarrow \quad \frac{\partial^2 \mathbf{V}}{\partial t^2} - c^2\,\Delta_3\ \boldsymbol{V} = \mu\,c^2\boldsymbol{J} \ , \ \frac{\partial^2 u}{\partial t^2} - c^2\,\Delta_3 u = \frac{c^2}{\epsilon}\rho$$

$$Div\mathbb{J} = 0 \quad \Leftrightarrow \quad \frac{\partial \rho}{\partial t} + div\,\boldsymbol{J} = 0 \qquad \text{(continuity equation)}$$

$$Div\,\mathbb{U} = 0 \qquad \Leftrightarrow \qquad \frac{\partial u}{\partial t} + c^2 div\,\mathrm{V} = 0 \qquad \text{(Lorentz condition)}$$

(Exercise 6). This four-dimensional formulation shows that the vectors (***E*,*B***) and (***D*,*H***) are coupled in pairs, the distinction between ***E*** and ***D***, ***B*** and ***H*** has an intrinsic character, and ***B*** is solenoidal in $\mathbb{R}^3$ for every t.

The field equations (5.22)−(5.27) and the transformation equations (5.17), (5.21) remain valid under any change of coordinates and for any system of general coordinates in the chronotope, provided the partial derivatives are replaced by tensorial derivatives. With this proviso the Maxwell equations are invariant not only with respect to Lorentz transformations, which correspond to passing from one inertial observer to another, but with respect to an arbitrary change of the space-time reference frame.

Exercises

Exercise 1. Eqs. (5.7), (5.8) imply

$$\begin{aligned} ds'^2 &= \alpha^2 dx_o^2 - \alpha^2\beta^2 dx_1^2 - 2i\alpha\beta dx_o dx_1 - \alpha^2\beta^2 dx_o^2 + \alpha^2 dx_1^2 + 2i\alpha\beta dx_o dx_1 + dx_2^2 + dx_3^2 \\ &= \alpha^2(1-\ \beta^2)\,dx_o^2 + \alpha^2(1-\beta^2)dx_1^2 + dx_2^2 + dx_3^2 \ = \ ds^2 \end{aligned}$$

since $\alpha^2(1-\beta^2) = 1$.

Exercise 2. Without loss of generality we may assume $\boldsymbol{v} = v\mathbf{c}_1$. The Maxwell equations in cartesian coordinates (x_1, x_2, x_3, t) are

$$\text{(E1)} \qquad div\, \boldsymbol{H} = 0 : \quad H_{1/x_1} + H_{2/x_2} + H_{3/x_3} = 0$$

$$\text{(E2)} \qquad div\, \boldsymbol{E} = 0 : \quad E_{1/x_1} + E_{2/x_2} + E_{3/x_3} = 0$$

$$\text{(E3)} \qquad curl\, \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \begin{cases} \mu\, H_{1/t} + E_{3/x_2} - E_{2/x_3} = 0 \\ \mu\, H_{2/t} + E_{1/x_3} - E_{3/x_1} = 0 \\ \mu\, H_{3/t} + E_{2/x_1} - E_{1/x_2} = 0 \end{cases}$$

$$\text{(E4)} \qquad curl \boldsymbol{H} = \epsilon \frac{\partial \boldsymbol{E}}{\partial t} \begin{cases} -\epsilon\, E_{1/t} + H_{3/x_2} - H_{2/x_3} = 0 \\ -\epsilon\, E_{2/t} + H_{1/x_3} - H_{3/x_1} = 0 \\ -\epsilon\, E_{3/t} + H_{2/x_1} - H_{1/x_2} = 0 \end{cases}$$

(the slashes denote partial derivatives). The Lorentz transformation (5.9) implies the relations between the first partial derivatives

$$\frac{\partial}{\partial x_1} = \alpha \frac{\partial}{\partial x_1'} - \alpha \frac{\beta}{c} \frac{\partial}{\partial t'} \;, \quad \frac{\partial}{\partial t} = \alpha \frac{\partial}{\partial t'} - \alpha v \frac{\partial}{\partial x_1'} \;, \quad \frac{\partial}{\partial x_2} = \frac{\partial}{\partial x_2'} \;, \quad \frac{\partial}{\partial x_3} = \frac{\partial}{\partial x_3'}$$

Substituting into the Maxwell equations, (E1) multiplied by α and the first eq. (E3) yield

$$\text{(E5)} \qquad \alpha^2 \frac{\partial H_1}{\partial x_1'} - \alpha^2 \frac{\beta}{c} \frac{\partial H_1}{\partial t'} + \frac{\partial\, \alpha H_2}{\partial x_2'} + \frac{\partial\, \alpha H_3}{\partial x_3'} = 0$$

$$\text{(E6)} \qquad \mu\, \alpha \frac{\partial H_1}{\partial t'} - \mu\, \alpha v \frac{\partial H_1}{\partial x_1'} + \frac{\partial E_3}{\partial x_2'} - \frac{\partial E_2}{\partial x_3'} = 0$$

Substituting the expression

$$-\alpha^2 \frac{\beta}{c} \frac{\partial H_1}{\partial t'} = -\frac{\beta}{c} (\alpha^2\, v \frac{\partial H_1}{\partial x_1'} - \frac{\alpha}{\mu} \frac{\partial E_3}{\partial x_2'} + \frac{\alpha}{\mu} \frac{\partial E_2}{\partial x_3'})$$

obtained from eq. (E6) into (E5) we obtain

$$\alpha^2 \frac{\partial H_1}{\partial x_1'} - \alpha^2 \frac{\beta v}{c} \frac{\partial H_1}{\partial x_1'} + \frac{\beta\alpha}{c\mu} \frac{\partial E_3}{\partial x_2'} - \frac{\beta\alpha}{c\mu} \frac{\partial E_2}{\partial x_3'} + \alpha \frac{\partial H_2}{\partial x_2'} + \alpha \frac{\partial H_3}{\partial x_3'} = 0$$

Since $\alpha^2 - \alpha^2\beta v/c = \alpha^2(1 - v^2/c^2) = 1$, this equation takes the same form as (E1)

$$div'\boldsymbol{H'} \equiv H'_{1/x'_1} + H'_{2/x'_2} + H'_{3/x'_3} = 0$$

if $\boldsymbol{H'}$ is defined by

$$H'_1 = H_1 \ , \ H'_2 = \alpha H_2 + \frac{\beta\alpha}{c\mu}E_3 \, , \ H'_3 = \alpha H_3 - \frac{\beta\alpha}{c\mu}E_2$$

This means that $\boldsymbol{H'} = \mu^{-1}\boldsymbol{B'}$ and $\boldsymbol{B'}$ is defined according to (5.13) by

$$B'_1 = B_1 \ , \ B'_2 = \alpha B_2 + \frac{\beta\alpha}{c}E_3 \, , \ B'_3 = \alpha B_3 - \frac{\beta\alpha}{c}E_2$$

Similarly, from (E2) multiplied by α and the first eq. (E4) we find

$$-\epsilon\alpha\frac{\partial E_1}{\partial t'} + \epsilon\alpha v\frac{\partial E_1}{\partial x'_1} + \frac{\partial H_3}{\partial x'_2} - \frac{\partial H_2}{\partial x'_3} = 0$$

whence

$$-\alpha^2\frac{\beta}{c}\frac{\partial E_1}{\partial t'} = \frac{\beta}{c}(-\alpha^2 v\frac{\partial E_1}{\partial x'_1} - \frac{\alpha}{\epsilon}\frac{\partial H_3}{\partial x'_2} + \frac{\alpha}{\epsilon}\frac{\partial H_2}{\partial x'_3})$$

Substituting into (E5) yields the equation

$$\frac{\partial E_1}{\partial x'_1} + \frac{\partial}{\partial x'_2}(\alpha E_2 - \frac{\beta\alpha}{c\epsilon}H_3) + \frac{\partial}{\partial x'_3}(\alpha E_3 + \frac{\beta\alpha}{c\epsilon}H_2) = 0$$

which has the same form as (E2)

$$div'\mathrm{E'} \equiv E'_{1/x'_1} + E'_{2/x'_2} + E'_{3/x'_3} = 0$$

if $\boldsymbol{E'}$ is defined by

(E7) $$E'_1 = E_1 \ , \ E'_2 = \alpha E_2 - \frac{\beta\alpha}{c\epsilon}H_3 \, , \ E'_3 = \alpha E_3 + \frac{\beta\alpha}{c\epsilon}H_2$$

Since $\boldsymbol{B} = \mu\boldsymbol{H}$ and

(E8) $$\frac{1}{c\epsilon} = c\mu$$

eq. (E7) coincides with the first row of (5.13).

The fact that c is the finite quantity satisfying (E8) implies that also the remaining Maxwell equations are form invariant. For example, the second eq. (E3)

$$\mu\frac{\partial H_2}{\partial t} + \frac{\partial E_1}{\partial x_3} - \frac{\partial E_3}{\partial x_1} = 0$$

becomes

$$\mu\frac{\partial H_2'}{\partial t'} + \frac{\partial E_1'}{\partial x_3'} - \frac{\partial E_3'}{\partial x_1'} + \frac{\beta\alpha}{\epsilon\, c}\frac{\partial H_2}{\partial x_1'} - \beta\alpha c\mu \ \frac{\partial H_2}{\partial x_1'} = 0 \qquad \text{(E9)}$$

and since, by force of (E8),

$$\frac{\beta\alpha}{\epsilon\, c} = \beta\alpha c\mu \equiv \alpha v\mu$$

the last two terms in eq. (E9) cancel out and we obtain the equation

$$\mu\frac{\partial H_2'}{\partial t'} + \frac{\partial E_1'}{\partial x_3'} - \frac{\partial E_3'}{\partial x_1'} = 0$$

which has the same form as the second eq. (E3). Finally the equation of continuity

$$\frac{\partial \rho}{\partial t} + div\, \boldsymbol{J} = 0 \qquad \text{(E10)}$$

becomes in the new coordinates

$$\alpha\frac{\partial \rho}{\partial t'} - \alpha v\frac{\partial \rho}{\partial x_1'} + \alpha\frac{\partial J_1}{\partial x_1'} - \frac{\beta\alpha}{c}\frac{\partial J_1}{\partial t'} + \frac{\partial J_2}{\partial x_2'} + \frac{\partial J_3}{\partial x_3'} = 0$$

that is

$$\frac{\partial}{\partial t'}(\alpha\rho - \frac{\beta\alpha}{c}J_1) + \frac{\partial}{\partial x_1'}(\alpha J_1 - \alpha v\rho) + \frac{\partial J_2}{\partial x_2'} + \frac{\partial J_3}{\partial x_3'} = 0$$

Thus eq. (E10) maintains the same form provided ρ' and $\boldsymbol{J}'$ are defined according to the third row of (5.13):

$$\rho' = \alpha\rho - \frac{\beta\alpha}{c}J_1\,,\ J_1' = \alpha J_1 - \alpha\beta c\rho \quad ,\ J_2' = J_2\,,\ J_3' = J_3$$

This discussion shows that the Maxwell equations are not invariant in form if $c \neq (\epsilon\mu)^{-\frac{1}{2}}$, and therefore are not invariant in form under Galilei transformations.

Exercise 3. For $\boldsymbol{J}= q\delta(\boldsymbol{x}-\boldsymbol{x}_q)\boldsymbol{v}$ eq. (1.10) yields

$$\boldsymbol{F}(\boldsymbol{x}_q) = \int_{\mathbb{R}^3} \boldsymbol{J}\wedge \boldsymbol{B}\, dV = q\int_{\mathbb{R}^3} \delta(\boldsymbol{x}-\boldsymbol{x}_q)\boldsymbol{v}\wedge \boldsymbol{B}\, dV = q\boldsymbol{v}\wedge \boldsymbol{B}|_{x=x_q}$$

and the Ampere rule yields the magnetic Lorentz force.

Exercise 4. Verify eqs. (5.13) and (5.20) using eqs. (5.16) and (5.19).

Exercise 5. Verify eqs. (5.21) .

Hint:

$$Rot\mathbb{U} = \begin{pmatrix} 0 & -\frac{i}{c}(u_{/x_1}+V_{1/t}) & -\frac{i}{c}(u_{/x_2}+V_{2/t}) & -\frac{i}{c}(u_{/x_3}+V_{3/t}) \\ \frac{i}{c}(u_{/x_1}+V_{1/t}) & 0 & V_{2/x_1}-V_{1/x_2} & V_{3/x_1}-V_{1/x_3} \\ \frac{i}{c}(u_{/x_2}+V_{2/t}) & V_{1/x_2}-V_{2/x_1} & 0 & V_{3/x_2}-V_{2/x_3} \\ \frac{i}{c}(u_{/x_3}+V_{3/t}) & V_{1/x_3}-V_{3/x_1} & V_{2/x_3}-V_{3/x_2} & 0 \end{pmatrix}$$

$$^*Rot\mathbb{U} = \begin{pmatrix} 0 & V_{3/x_2}-V_{2/x_3} & V_{1/x_3}-V_{3/x_1} & V_{2/x_1}-V_{1/x_2} \\ V_{2/x_3}-V_{3/x_2} & 0 & -\frac{i}{c}(u_{/x_3}+V_{3/t}) & \frac{i}{c}(u_{/x_2}+V_{2/t}) \\ V_{3/x_1}-V_{1/x_3} & \frac{i}{c}(u_{/x_3}+V_{3/t}) & 0 & -\frac{i}{c}(u_{/x_1}+V_{1/t}) \\ V_{1/x_2}-V_{2/x_1} & -\frac{i}{c}(u_{/x_2}+V_{2/t}) & \frac{i}{c}(u_{/x_1}+V_{1/t}) & 0 \end{pmatrix}$$

$(j,k=0,...,3)$, so that

$$(\mathbf{Div}\,^*Rot\mathbb{U})_o = V_{3/x_2x_1} - V_{2/x_3x_1} - V_{3/x_1x_2} + V_{1/x_3x_2} + V_{2/x_1x_3} - V_{1/x_2x_3} = 0$$

$$(\mathbf{Div}\,^*Rot\mathbb{U})_1 = V_{2/x_3x_0} - V_{3/x_2x_0} - \frac{i}{c}u_{/x_3x_2} + V_{3/x_0x_2} + \frac{i}{c}u_{/x_2x_3} - V_{2/x_ox_3} = 0$$

$$(\mathbf{Div}\,^*Rot\mathbb{U})_2 = V_{3/x_1x_0} - V_{1/x_3x_0} + \frac{i}{c}u_{/x_3x_1} - V_{3/x_0x_1} - \frac{i}{c}u_{/x_1x_3} + V_{1/x_ox_3} = 0$$

$$(\mathbf{Div}\,^*Rot\mathbb{U})_3 = V_{1/x_2x_0} - V_{2/x_1x_0} - \frac{i}{c}u_{/x_2x_1} + V_{2/x_0x_1} + \frac{i}{c}u_{/x_1x_2} - V_{1/x_ox_2} = 0$$

Exercise 6. Verify the correspondence table given in the text.

Chapter 6

Anisotropy, Dispersion and Nonlinearity

The linear local constitutive relations (C1), (C2) considered in Chapter 1

(C1) $$\boldsymbol{J} = \gamma \boldsymbol{E}$$

(C2) $$\boldsymbol{D} = \epsilon \boldsymbol{E}$$

are to be viewed as an approximate model of physical reality:real media, whether conducting or not, are dispersive and nonlinear. Dispersive means that the polarization vector $\boldsymbol{P}$ and the current density $\boldsymbol{J}$ depend also on past values of $\boldsymbol{E}$ via a nonlocal hereditary relation. For a homogeneous isotropic dielectric we may write

(6.1) $$\boldsymbol{P}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\psi}(\tau) \boldsymbol{E}(\boldsymbol{x},\, t-\tau) d\tau$$

where the real-valued function $\tilde{\psi}(\tau)$, called the memory function for the polarization, satisfies the causality relation

(6.2) $$\tilde{\psi}(\tau) \equiv 0 \quad \text{for all} \quad \tau < 0$$

and depends on the material under consideration. Since $\boldsymbol{D} = \epsilon_o \boldsymbol{E} + \boldsymbol{P}$ (see §1.8), eq. (6.1) shows that (C2) should be replaced by the nonlocal relation

(6.3) $$\boldsymbol{D}(\boldsymbol{x},\, t) = \epsilon_o \boldsymbol{E}(\boldsymbol{x},\, t) + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\psi}(\tau) \boldsymbol{E}(\boldsymbol{x},\, t-\tau) d\tau$$

which has the same form as the equation

$$\boldsymbol{D}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\epsilon}(\tau)\boldsymbol{E}(\boldsymbol{x},\, t-\tau)d\tau \tag{6.4}$$

written for a homogeneous conductor in Chapter 1, if we set

$$\tilde{\epsilon}(\tau) = \sqrt{2\pi}\epsilon_o\delta(\tau) + \tilde{\psi}(\tau) \tag{6.5}$$

where $\delta(\tau)$ is the one-dimensional Dirac distribution (Exercise 1). Thus for a generic dielectric medium the memory function $\tilde{\epsilon}(\tau)$ for $\boldsymbol{D}$ will be a distribution.

Similarly, Ohm's law (C1) for a homogeneous isotropic conductor should be written in the hereditary form

$$\boldsymbol{J}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\gamma}(\tau)\boldsymbol{E}(\boldsymbol{x},\, t-\tau)d\tau \tag{6.6}$$

The real-valued function $\tilde{\gamma}(\tau)$, called the memory function for the current, satisfies the causality relation

$$\tilde{\gamma}(\tau) \equiv 0 \quad \text{for all} \quad \tau < 0 \tag{6.7}$$

and depends upon the material, while the polarization $\tilde{\psi}(\tau)$ for a good conductor is small. In a (good) conductor the displacement current is negligible with respect to the conduction current. Inside a perfect conductor we have simply $\boldsymbol{E} \equiv \boldsymbol{D} \equiv \boldsymbol{0}$ (see Proposition 1.3.2): this case is trivial and will be excluded in the sequel.

For time-periodic phenomena, the convolution theorem implies that the Fourier transforms of the permittivity and conductivity memory functions $\tilde{\epsilon}(\tau)$, $\tilde{\gamma}(\tau)$ depend on the frequency (§1). However, we will see that if one is interested to a narrow frequency range outside the regions of anomalous dispersion, the permittivity and the conductivity can in practice be viewed as frequency-independent and the local relations (C1) and (C2) can be retained, with suitable values for γ and ϵ (see §6.2 and §6.3). This is all the more true for the magnetic permeability, and for this reason the linear local relation (C3) of Chapter 1 with $\mu = \mu_o$

$$\boldsymbol{B} = \mu_o\boldsymbol{H} \tag{6.8}$$

will be applied to all non-magnetic materials thoughout this chapter.

In anisotropic media, such as crystals, the permittivity and the conductivity, i.e. the scalar memory functions $\tilde{\psi}(\tau)$, $\tilde{\epsilon}(\tau)$ and $\tilde{\gamma}(\tau)$, are replaced by real symmetric rank-2 tensors, for which the same considerations apply (see §6.4 and §6.5).

In addition, real media are nonlinear, although this fact can usually be neglected in practice. The conductivity tensor may depend on the magnetic field strength, giving rise to the nonlinear Hall effect (§6.6). More to the point here, in the presence of very strong electric fields the permittivity can depend on the strength of the electric field so that the constitutive relation $\boldsymbol{D}$=$\boldsymbol{D(E)}$ becomes nonlinear:this gives rise to a number of nonlinear effects, like the generation of optical harmonics and the Kerr effect [12,35]. We will discuss here a mathematical model, based on the nonlinear Maxwell equations, for the phenomenon of second-harmonic generation in nonlinear optics (§6.7).

6.1 Hereditary constitutive relations

6.1.1 Memory functions.

Consider a homogeneous isotropic medium and a time-harmonic electric field given by the (real part of the) expression

$$\boldsymbol{E}(\boldsymbol{x},t)=\mathbb{E}(\boldsymbol{x})e^{i\omega t} \tag{6.9}$$

where $\mathbb{E}(\boldsymbol{x})$ is a real-valued function of $\boldsymbol{x}$ (which may depend also on ω, see §6.3). Substituting this expression in eq. (6.3) we obtain

$$\begin{aligned}\boldsymbol{D}(\boldsymbol{x},\ t)&=\epsilon_o\mathbb{E}(\boldsymbol{x})e^{i\omega t}+\mathbb{E}(\boldsymbol{x})e^{i\omega t}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\tilde{\psi}(\tau)e^{-i\omega\tau}d\tau\\&=(\epsilon_o+\psi(\omega))\ \boldsymbol{E}(\boldsymbol{x},t)\end{aligned}$$

where

$$\psi(\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-i\omega\tau}\tilde{\psi}(\tau)d\tau \qquad (\omega\in\mathbb{R}) \tag{6.10}$$

is the Fourier transform of the distribution $\tilde{\psi}(\tau)$. Equivalently, if we start from eq. (6.4) we obtain

$$\boldsymbol{D}(\boldsymbol{x},\, t) = \epsilon(\omega)\boldsymbol{E}(\boldsymbol{x}, t) \equiv \epsilon(\omega)\mathbb{E}(\boldsymbol{x})e^{i\omega t} \tag{6.11}$$

where

$$\epsilon(\omega) = \epsilon_o + \psi(\omega) \tag{6.12}$$

is the Fourier transform of the memory function (distribution) $\tilde{\epsilon}(\tau)$ given by eq. (6.5). Note that in real notations eq. (6.11) reads

$$\boldsymbol{D}(\boldsymbol{x},\, t) = \big[(\epsilon_o + Re\tilde{\psi}(\omega))\, cos\omega t - Im\tilde{\psi}(\omega) sin\omega t\big]\mathbb{E}(\boldsymbol{x})$$

Similarly, eq. (6.6) implies that the current density generated by the electric field $\mathbb{E}(\boldsymbol{x})e^{i\omega t}$ in a conductor is

$$\boldsymbol{J}(\boldsymbol{x},\, t) = \gamma(\omega)\boldsymbol{E}(\boldsymbol{x},\, t) \equiv \gamma(\omega)\mathbb{E}(\boldsymbol{x})e^{i\omega t} \tag{6.13}$$

where

$$\gamma(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-i\omega\tau}\tilde{\gamma}(\tau)d\tau \qquad (\omega \in \mathbb{R}) \tag{6.14}$$

is the Fourier transform of the memory function $\tilde{\gamma}(\tau)$. The assumptions H1−H3 below, which follow from experimental results [33] and from theoretical considerations [35], are essential in view of the physical interpretation of $\epsilon(\omega)$ and $\gamma(\omega)$.

H1. The memory functions $\tilde{\psi}(t)$, $\tilde{\gamma}(t)$ are bounded and continuous for $t \geq 0$ and admit Fourier transforms $\psi(\omega)$, $\gamma(\omega)$ ($\omega \in \mathbb{R}$), vanishing at infinity and satisfying the inversion formulae

$$\tilde{\psi}(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega t}\psi(\omega)\, d\omega \tag{6.15}$$

$$\tilde{\gamma}(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega t}\gamma(\omega)\, d\omega \quad (t \in \mathbb{R})$$

in the sense of distributions.

The causality relations (6.2), (6.7) imply that the (distributional) Fourier transforms

$$(6.16)\quad \psi(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-i\omega\tau} \tilde{\psi}(\tau) d\tau \quad , \quad \gamma(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-i\omega\tau} \tilde{\gamma}(\tau) d\tau$$

($\omega \in \mathbb{R}$) have real parts

$$Re\psi(x) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} cos(x\tau)\, \tilde{\psi}(\tau) d\tau \quad , \quad Re\gamma(x) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} cos(x\tau)\, \tilde{\gamma}(\tau) d\tau$$

that are even functions of $x = \omega \in \mathbb{R}$, and imaginary parts

$$Im\psi(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} sin(x\tau)\, \tilde{\psi}(\tau) d\tau \ , \ Im\gamma(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} sin(x\tau)\, \tilde{\gamma}(\tau) d\tau$$

that are odd functions of $x = \omega \in \mathbb{R}$. In particular, the static permittivity

$$(6.17)\qquad \epsilon(0) = \epsilon_o + \psi(0) \ = \epsilon_o + \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \tilde{\psi}(\tau)\, d\tau$$

and the static conductivity for a conductor

$$\gamma(0) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \tilde{\gamma}(\tau)\, d\tau$$

are always real. They must also be positive. For example, $\epsilon(0) \cong 80\epsilon_o$, $\gamma(0) \cong 0$ for distilled water. Moreover, assumption **H1** implies that $\epsilon(\omega)$ approaches ϵ_o as $\omega \to \infty$.

Eqs. (6.11), (6.12) and (6.13) show that the Fourier transforms $\epsilon(\omega)$, $\psi(\omega)$, $\gamma(\omega)$, called complex permittivity, complex polarizability and complex conductivity at the frequency ω, respectively, can be interpreted as the permittivity and conductivity in the frequency space, i.e. for monochromatic waves. We will show that they are necessarily complex-valued for finite $\omega \neq 0$.

The next assumption implies a sort of causality principle.

H2. The Fourier transforms $\psi(\omega)$ and $\gamma(\omega)$, $\omega \in \mathbb{R}$, can be extended to the complex plane $\omega \in \mathbb{C}$ as holomorphic functions in the lower-half plane, vanishing at infinity and with a finite number of pole singularities in the upper-half plane $Im\omega > 0$.

In other words, $\psi(\omega)$ and $\gamma(\omega)$ are holomorphic for $Im\omega < 0$ and can be continued analytically for $Im\omega \geq 0$ with the exception of a finite number of poles having positive imaginary part [1]. The next assumption can be envisaged as a definition.

H3. For a good conductor $\psi(\omega) \cong 0$ and $\epsilon(\omega) \cong \epsilon_o$, for a bad conductor $\gamma(\omega) \cong 0$.

Since $\psi(\omega)$ and $\gamma(\omega)$ have no singularities on the real axis the Hilbert integral formulae [38] hold for $Im\ \omega < 0$:

$$\psi(\omega) = \frac{1}{\pi i} \int_{-\infty}^{+\infty} \frac{Re\psi(x)}{\omega - x} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{Im\psi(x)}{\omega - x} dx \tag{6.18}$$

$$\gamma(\omega) = \frac{1}{\pi i} \int_{-\infty}^{+\infty} \frac{Re\gamma(x)}{\omega - x} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{Im\gamma(x)}{\omega - x} dx \tag{6.19}$$

The analytic continuations of these functions are holomorphic in the half−plane $Im\,\omega < k$, where $k > 0$ is the smallest imaginary part of the poles. An immediate consequence of this fact is the following

Proposition 6.1.1 *$Re\psi(\omega)$ and $Im\psi(\omega)$ can have at most isolated zeroes for $\omega \in \mathbb{R}$. Similarly for $\gamma(\omega)$.*

Letting $Im\,\omega \to -0$ in eq. (6.18), the Plemelyi formulae yield

$$\psi(\omega) = Re\psi(\omega) + \frac{1}{\pi i} P \int_{-\infty}^{+\infty} \frac{Re\psi(x)}{\omega - x} dx \qquad (\omega \in \mathbb{R})$$

or

$$\psi(\omega) = iIm\psi(\omega) + \frac{1}{\pi}\ P \int_{-\infty}^{+\infty} \frac{Im\psi(x)}{\omega - x} dx \qquad (\omega \in \mathbb{R})$$

where P denotes the Cauchy principal value of the integral [38]. Separating the real and imaginary parts we obtain the Kramers-Kronig relations for the Fourier transform of the complex polarizability $\psi(\omega)$

$$Re\,\psi(\omega\ ) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{Im\psi(x)}{\omega - x} dx \qquad (\omega \in \mathbb{R}) \tag{6.20}$$

[1]For a time-dependence of the form $exp(-i\omega t)$ the role of the upper and lower half-planes is exchanged and other sign modifications in the formulae of this section are required

(6.21) $$Im\,\psi(\omega\,) = -\frac{1}{\pi}\;P\int_{-\infty}^{+\infty}\frac{Re\psi(x)}{\omega - x}dx \qquad (\omega\in\mathbb{R})$$

In the case of a conductor, similar relations hold for the complex conductivity $\gamma(\omega)$.

Proposition 6.1.2 *The memory functions $\tilde{\psi}(t)$, $\tilde{\gamma}(t)$ vanish as $t \to +\infty$ (fading memories).*

Proof. $\psi(\omega)$ has a finite number of poles $\omega_1\,, ...,\, \omega_p$ with $Im\omega_k > 0$ for all $k = 1, ..., p$. Moreover $\psi(\omega)$ vanishes at infinity and is holomorphic for $Im\omega < 0$. The Jordan Lemma (see [45] p. 303) then implies that

$$\lim_{R\to\infty}\int_{\Gamma_R} e^{i\omega t}\psi(\omega)d\omega = 0$$

where Γ_R is a semicircle of radius R centered at the origin in the upper-half plane for $t > 0$, in the lower-half plane for $t < 0$. Thus the path of integration in the integral (6.15) for $\tilde{\psi}(t)$ can be made into a closed path by adding the semicircle Γ_R, and applying the theorem of residues gives

$$\tilde{\psi}(t) \equiv 0 \qquad \text{for } t < 0$$

and

(6.22) $$\tilde{\psi}(t) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega t}\;\psi(\omega)d\omega = \sqrt{2\pi}\;i\;\sum_{k=1}^{p} e^{i\omega_k t}\,\mathfrak{R}_k \qquad \text{for } t > 0$$

where $\mathfrak{R}_k$ is the residue of $\psi(\omega)$ at ω_k. Since

$$e^{i\omega_k t} = e^{-Im\omega_k t}\,e^{iRe\omega_k t} \;\to 0 \qquad \text{as } t \to +\infty$$

the conclusion follows. The proof for $\gamma(\omega)$ is entirely similar.

Eq. (6.5) and Proposition 6.1.2 imply that

(6.23) $$\lim_{t\to\infty}\tilde{\epsilon}(t) = 0$$

Proposition 6.1.3 *Suppose that a memory function $\tilde{\psi}(t)$ (or $\tilde{\gamma}(t)$) satisfies the following assumptions*:

(*i*) $\tilde{\psi}(t) \in C^2[0, +\infty)$, $\tilde{\psi}(t) \equiv 0$ *for* $t < 0$,

(*ii*) $\tilde{\psi}(t)$ *is a positive convex decreasing function of* t *for* $t > 0$ $(\tilde{\psi} > 0, \dot{\tilde{\psi}} < 0, \ddot{\tilde{\psi}} > 0)$

(*iii*) $\lim_{t\to\infty} \tilde{\psi}(t) = 0$ *(fading memory).*

Then $Re\psi(\omega) > 0$ *for all* ω.

Proof. See Theorem 124, p. 170 of E.C. Titchmarsh, Introduction to the Theory of Fourier Integrals, Oxford University Press, 1948.

This result implies that an exponential memory function of the type $\tilde{\psi}(t) = \psi_o e^{-kt}$ $(k > 0)$ is not possible if $Re\psi(\omega)$ changes sign for $\omega > 0$. As we will see in §6.2, the experiments show that this is indeed the case for a dielectric or a bad conductor.

Remark 1. Since $Re\psi(\omega) = Re\psi(-\omega), Im\psi(\omega) = -Im\psi(-\omega)$, the memory functions (6.15) can be written as

$$(6.24)\quad \tilde{\epsilon}(t) = \epsilon_o + \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{+\infty} Re\psi(\omega) cos\,\omega t\; d\omega + i\frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{+\infty} Im\psi(\omega) sin\,\omega t\, d\omega$$

$$(6.25)\quad \tilde{\gamma}(t) = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{+\infty} Re\gamma(\omega)\;\; cos\,\omega t\, d\omega + i\frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{+\infty} Im\gamma(\omega)\, sin\,\omega t\; d\omega$$

Thus only nonnegative values of $\omega \in \mathbb{R}$ need to be considered.

Remark 2. To summarize:

(*i*) the static polarization $\psi(0)$ and the static permittivity $\epsilon(0)$ are positive and finite

(*ii*) for a conductor the static conductivity $\gamma(0)$ is positive and finite

(*iii*) the Fourier transforms $\psi(\omega)$, $\gamma(\omega)$ are analytic in a half−plane $Im\,\omega < k$, $k > 0$.

Remark 3 (non-dispersive case). If $\tilde{\psi}$ is the Dirac distribution

$$(6.26)\qquad \tilde{\psi}(\tau) = \sqrt{2\pi}\; (\epsilon - \epsilon_o)\, \delta(\tau)\;\;, \qquad \tilde{\epsilon}(\tau) = \sqrt{2\pi}\, \epsilon\, \delta(\tau)$$

then $\psi(\omega) = \epsilon - \epsilon_o$, so that

$$Im\, \psi(\omega) = 0$$

and eq. (6.4) reduces to the local constitutive relation

$$\boldsymbol{D} = \epsilon \boldsymbol{E} \tag{6.27}$$

(see § 6.3). Viceversa, (6.27) implies (6.26). Similarly if

$$\tilde{\gamma}(\tau) = \sqrt{2\pi}\gamma\, \delta(\tau) \tag{6.28}$$

then $\gamma(\omega) = \gamma$ and eq. (6.6) reduces to Ohm's law

$$\boldsymbol{J} = \gamma\, \boldsymbol{E}$$

Conversely, Ohm's law implies (6.28). In this limit case assumptions **H1** and **H2** are not satisfied.

6.1.2 Dispersion relations.

The imaginary part of $\psi(\omega)$ is related to the physical phenomen of absorption. In order to see this we need to find the form of the dispersion relation in the hereditary case. As in §4.2, we try to determine the wavenumber $\boldsymbol{p}$=$\boldsymbol{p}(\omega)$ in such a way that the plane monochromatic wave [2]

$$\boldsymbol{E} = \boldsymbol{E}_o e^{i(\omega t - \,\boldsymbol{p}(\omega)\cdot\boldsymbol{x})} \quad , \quad \boldsymbol{H} = \boldsymbol{H}_o e^{i(\omega t - \,\boldsymbol{p}(\omega)\cdot\boldsymbol{x})}$$

satisfies for a given frequency ω the Maxwell equations

$$\frac{\partial \boldsymbol{D}}{\partial t} = curl\, \boldsymbol{H} - \boldsymbol{J} \;\;,\;\; \mu_o \frac{\partial \mathbf{H}}{\partial t} = -curl\, \mathbf{E} \;,\; div\, \boldsymbol{E} \; = div\, \boldsymbol{H} \; = 0$$

with $\boldsymbol{D}$, $\boldsymbol{J}$ given by the nonlocal constitutive relations (6.4), (6.6). Since here $\mathbb{E}(\boldsymbol{x})$ =$\boldsymbol{E}_o e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}$, eqs. (6.11) and (6.13) become

$$\boldsymbol{D}(\boldsymbol{x},\, t) = \epsilon(\omega)\boldsymbol{E}_o e^{i(\omega t - \,\boldsymbol{p}(\omega)\cdot\boldsymbol{x})} \;,\; \boldsymbol{J}(\boldsymbol{x},\, t) = \gamma(\omega)\boldsymbol{E}_o e^{i(\omega t - \,\boldsymbol{p}(\omega)\cdot\boldsymbol{x})}$$

---

[2] we recall that this is understood to mean that $\boldsymbol{E}$= $Re(\boldsymbol{E}_o e^{i\omega t - i\,\boldsymbol{p}(\omega)\cdot\boldsymbol{x}})$, $\boldsymbol{H}$= $Re(\boldsymbol{H}_o e^{i\omega t - i\,\boldsymbol{p}(\omega)\cdot\boldsymbol{x}})$, and $\boldsymbol{S}$= $Re(\boldsymbol{E}_o e^{i\omega t - i\,\boldsymbol{p}(\omega)\cdot\boldsymbol{x}}) \wedge Re(\boldsymbol{H}_o e^{i\omega t - i\,\boldsymbol{p}(\omega)\cdot\boldsymbol{x}})$

Proceeding as in §4.2 we find the algebraic equations

$$\boldsymbol{p}\cdot\boldsymbol{E}_o = 0\ ,\ \ \boldsymbol{p}\cdot\boldsymbol{H}_o = 0\ ,\ \ i\boldsymbol{p}\wedge\boldsymbol{E}_o = i\mu_o\omega\boldsymbol{H}_o\ ,\ \ -i\boldsymbol{p}\wedge\boldsymbol{H}_o = i\omega\epsilon'(\omega)\,\boldsymbol{E}_o$$

where

$$\epsilon'(\omega) := \epsilon(\omega) - i\frac{\gamma(\omega)}{\omega} \ \equiv\ \epsilon_o + \psi(\omega) - i\frac{\gamma(\omega)}{\omega} \tag{6.29}$$

is the counterpart of the complex permittivity of the conductor at the frequency ω introduced in eq. (4.12). Since $\gamma(0) > 0$ for a conductor, eq. (6.29) immediately implies the following result.

Proposition 6.1.4 *For a conductor the imaginary part of $\epsilon'(\omega)$ has a pole at $\omega = 0$.*

Upon elimination of $\boldsymbol{H}$ we find the dispersion relation for a conductor

$$\boldsymbol{p}\cdot\boldsymbol{p} \ = \ \frac{\epsilon'(\omega)}{\epsilon_o}\frac{\omega^2}{c_o^2} \tag{6.30}$$

which holds also for a dielectric if we take $\gamma(\omega) \equiv 0$ and $\epsilon'(\omega) \equiv \epsilon(\omega)$. In other words, the wavenumber $\boldsymbol{p}{=}\boldsymbol{p}(\omega)$ must satisfy

$$\boldsymbol{p}\cdot\boldsymbol{p} \ = \ \frac{1}{c_o^2}\left[\omega^2 + \frac{\omega^2}{\epsilon_o}\psi(\omega) - i\frac{\omega\gamma(\omega)}{\epsilon_o}\right] \tag{6.31}$$

for a conductor, and

$$\boldsymbol{p}\cdot\boldsymbol{p} \ = \ \frac{\epsilon(\omega)}{\epsilon}\frac{\omega^2}{c_o^2} \equiv \left[1 + \frac{1}{\epsilon_o}\psi(\omega)\right]\frac{\omega^2}{c_o^2} \tag{6.32}$$

for a dielectric.

Thermodynamic considerations [17,20] also show that the following assumption must be satisfied.

H4. $\epsilon'(\omega)$ *must satisfy the inequality* $\omega\, Im\epsilon'(\omega) \le 0$ for all $\omega \in \mathbb{R}$, i.e.

$$\omega\, Im\epsilon(\omega) \le Re\gamma(\omega)$$

This assumption is satisfied in particular if $\omega\, Im\epsilon(\omega) < 0$, $Re\gamma(\omega) > 0$.

Proposition 6.1.5 *$Imp(\omega)$ cannot be identically zero for $\omega \in \mathbb{R}$.*

Proof. $\boldsymbol{p}(\omega)\cdot\boldsymbol{p}(\omega)$ is analytic in a neighborhood of $\omega = 0$ by force of **H2**, and its imaginary part has at most isolated zeroes by force Proposition 6.1.1. Hence $Im\boldsymbol{p}(\omega)$ cannot be identically zero, even if $\gamma(\omega) \equiv 0$.

Thus in all cases the wavenumber $\boldsymbol{p}$=$\boldsymbol{p}(\omega)$ must have the form

$$\boldsymbol{p} = \boldsymbol{P}' - i\boldsymbol{p}' \tag{6.33}$$

where $\boldsymbol{p}', \boldsymbol{P}' \in \mathbb{R}^3\backslash\{\boldsymbol{0}\}$ satisfy the dispersion relation (6.31) or (6.32), that is

$$\boldsymbol{p}\cdot\boldsymbol{p} \equiv |\boldsymbol{P}'|^2 - |\boldsymbol{p}'|^2 - 2i\boldsymbol{P}'\cdot\boldsymbol{p}' = \frac{\epsilon'(\omega)}{\epsilon_o}\frac{\omega^2}{c_o^2}$$

(In the case of a dielectric, $\epsilon'(\omega)$ must be replaced by $\epsilon(\omega)$.) It follows that the plane wave

$$\boldsymbol{E} = \boldsymbol{E}_o e^{-\boldsymbol{p}'(\omega)\cdot\boldsymbol{x}}\, e^{i(\omega t - \boldsymbol{P}'(\omega)\cdot\boldsymbol{x})}, \qquad \boldsymbol{H} = \boldsymbol{H}_o e^{-\boldsymbol{p}'(\omega)\cdot\boldsymbol{x}}\, e^{i(\omega t - \boldsymbol{P}'(\omega)\cdot\boldsymbol{x})}$$

contains in any case a damping term $e^{-\boldsymbol{p}'(\omega)\cdot\boldsymbol{x}}$, typical of absorbing media (cfr. §4.2 of). We have proven the following

Proposition 6.1.6 *Under assumptions H1−H3 all dispersive media (conductors or dielectrics) are absorbing media.*

6.2 Evaluation of the memory functions

Very little is known about the memory functions $\tilde{\psi}(\tau)$ and $\tilde{\gamma}(\tau)$, with the exception of their overall behavior and a few values of their Fourier transforms $\psi(\omega)$, $\gamma(\omega)$. Some insight may be gained by means of a simple mechanical model of the electron based on the equation of forced oscillations [42]. Let

$$\boldsymbol{x} = \boldsymbol{x}_o e^{i\omega t} \tag{6.34}$$

denote the displacement from its equilibrium position of an electron acted upon by the force $-e\boldsymbol{E}$ due to a monochromatic electric field of the form

$$\boldsymbol{E} = \boldsymbol{E}_o e^{i\omega t}$$

Consider first a dielectric. In this case the electrons are bound to the nuclei and we may assume that there is a linear restoring force $-\omega_o^2\boldsymbol{x}$ proportional to the displacement $\boldsymbol{x}$, as well as a damping force $-g\boldsymbol{v}$ proportional to the velocity $\boldsymbol{v}= d\boldsymbol{x}/dt$, to account for absorption (cfr. Proposition 6.1.6 above). The equation of motion for the electron is then

$$\frac{d^2\boldsymbol{x}}{dt^2} + g\frac{d\boldsymbol{x}}{dt} + \omega_o^2\boldsymbol{x} = -\frac{e}{m}\boldsymbol{E}_o e^{i\omega t}$$

where m is the electron mass. For a solution $\boldsymbol{x}$ of the form (6.34) we find $\boldsymbol{v}= i\omega\boldsymbol{x}$ and

$$-\omega^2\boldsymbol{x}_o + i\omega g\boldsymbol{x}_o + \omega_o^2\boldsymbol{x}_o = -\frac{e}{m}\boldsymbol{E}_o$$

or

$$\boldsymbol{x}_o = \frac{e}{m}\frac{1}{\omega^2 - \omega_o^2 - i\omega g}\boldsymbol{E}_o \tag{6.35}$$

The polarization vector $\boldsymbol{P}\equiv (\tilde{\epsilon}(\omega) - \epsilon_o)\boldsymbol{E}_o e^{i\omega t}$ at the frequency ω will then be given by

$$\boldsymbol{P} = -Ne\boldsymbol{x} \equiv -Ne\boldsymbol{x}_o e^{i\omega t}$$

so that

$$\psi(\omega)\boldsymbol{E}_o \equiv (\epsilon(\omega) - \epsilon_o)\boldsymbol{E}_o = -Ne\boldsymbol{x}_o$$

where N is the number of electrons per unit volume of the material under consideration, and $-e < 0$ is the electron charge. From eq. (6.35) we find the expression

$$\epsilon(\omega) = \epsilon_o - \frac{Ne^2}{m}\frac{1}{\omega^2 - \omega_o^2 - i\omega g}$$

for the complex permittivity at the frequency ω or, since $\epsilon(\omega) = \epsilon_o + \psi(\omega)$ (eq. (6.12)), the expression

$$\psi(\omega) = -\frac{Ne^2}{m}\frac{1}{\omega^2 - \omega_o^2 - i\omega g} \tag{6.36}$$

for the complex polarizability at the frequency ω. It follows that

$$Re\psi(\omega) = \frac{Ne^2}{m}\frac{\omega_o^2 - \omega^2}{(\omega^2 - \omega_o^2)^2 + \omega^2 g^2} \quad , \quad Im\psi(\omega) = -\frac{Ne^2}{m}\frac{\omega\, g}{(\omega^2 - \omega_o^2)^2 + \omega^2 g^2}$$

so that the static polarizability and the static permittivity

$$\psi(0) = \frac{Ne^2}{m\omega_o^2}\,, \qquad \epsilon(0) = \epsilon_o + \frac{Ne^2}{m\omega_o^2} \tag{6.37}$$

are real and positive. Since $\omega^2 - \omega_o^2 - i\omega g = 0$ has the two roots

$$\omega = \omega_\pm := i(\frac{g}{2} \pm \frac{1}{2}\sqrt{g^2 - 4\omega_o^2}\,)$$

we see that for $g \neq 2\omega_o$, $\psi(\omega)$ has two simple poles $\omega_\pm$ with positive imaginary part in the complex plane ω, and is holomorphic otherwise. For $g > 2\omega_o$ the two poles lie on the imaginary axis, and for $g = 2\omega_o$ they coalesce into a single pole of order two at $\omega_+ = \omega_- = ig/2$. Thus **H2** is satisfied. Moreover, $Im\psi(\omega)$ never vanishes and is always negative, while $Re\psi(\omega)$ has a zero at $\omega = \omega_o$ $(\omega = -\omega_o)$ and changes sign at $\omega = \omega_o$ (cfr. Proposition 6.1.1) .

Calculating the inverse Fourier transform of $\psi(\omega)$ in eq. (6.36) (Exercise 2) yields the expression for the polarization memory function for $g \neq 2\omega_o$
(6.38)

$$\tilde{\psi}(t) = \sqrt{8\pi}\,\frac{Ne^2}{m}\frac{sinh(\sqrt{g^2-4\omega_o^2}\,t/2)}{\sqrt{g^2-4\omega_o^2}}\,e^{-gt/2}\ (t \geq 0)\,;\ \tilde{\psi}(t) \equiv 0\,(t<0)$$

For $g < 2\omega_o$ $\tilde{\psi}(t)$ has an oscillating behavior with an infinite number of zeroes for $t > 0$. If $g = 2\omega_o$ we find, by direct calculation or by passing to the limit in eq. (6.38),

$$\tilde{\psi}(t) = \sqrt{2\pi}\,\frac{Ne^2}{m}t\,e^{-gt/2}\ \ (t \geq 0)\ \ ;\ \ \tilde{\psi}(t) \equiv 0\ \ (t<0) \tag{6.39}$$

In all cases **H1** is satisfied. The numerical values of the quantities g and ω_o can be estimated as follows: N is of order 10^{29}, m of order $10^{-30}\,kg$, e of order $1.6\,10^{-19}coulomb$, ϵ_o of order $10^{-11}farad/m$ and the static polarizability $\psi(0) \cong 1.6\epsilon_o$, so that eq. (6.36) gives the estimate for the zero of $Re\psi(\omega)$

$$\omega_o \cong \sqrt{\frac{Ne^2}{m\epsilon_o}} \simeq \sqrt{\frac{10^{29}1.6\,10^{-38}}{10^{-30}1.6\,10^{-11}}} \cong\ 10^{16}\,sec^{-1}$$

This value ω_o for the frequency corresponds to a wavelength in vacuo

$$\lambda_o = \frac{2\pi}{\omega_o}\, c_o \cong \frac{6}{10^{16}}\, 3\, 10^8 \cong 1.8\, 10^{-7} m \; = 1800\, \text{Å} \tag{6.40}$$

(cfr. [22], p. 334). If we assume that $g > 2\omega_o$, the function $Re\psi(\omega)$ turns out to be decreasing for $0 < \omega < \sqrt{\omega_o(\omega_o + g)}$ (anomalous dispersion) and increasing for $\omega > \sqrt{\omega_o(\omega_o + g)}$ (normal dispersion) (Exercise 3). The dispersion is called anomalous if $Re\psi(\omega)$ decreases for decreasing wavelength (increasing frequency) and $|Im\psi(\omega)|$ is high, giving rise to high absorption. In contrast, normal dispersion is characterized by the fact that $Re\psi(\omega)$ increases for decreasing wavelength and $|Im\psi(\omega)|$ is negligible.

This evaluation of the memory function, however, accounts only for the so-called "ultraviolet anomalous dispersion" due to the electrons' polarizability, and the behavior of $\psi(\omega)$ derived above is correct only for high frequencies, $\omega > \omega_o$. For lower frequencies $\omega < \omega_o$ (higher wavelengths $\lambda > \lambda_o$) one must take into account also contributions due to atomic and dipolar polarization, and two further frequency intervals of anomalous dispersion (absorption bands) arise at higher wavelengths, i.e. lower frequencies. The corresponding graph of the function $\epsilon(\omega)$ for $\omega > 0$, derived from experiments [33], is reported in Fig. 6.1. (We recall that $Re\psi(\omega)$ is even in ω while $Im\psi(\omega)$ is odd.)

The figure shows three intervals of anomalous dispersion or absorption bands $(\omega_1,\, \omega_2)$, (ω_3, ω_4), $(\omega_5,\, \omega_6)$ due to the dipolar, atomic and electronic contributions, respectively, and three frequency intervals where $Re\ \epsilon(\omega) = \epsilon_o + Re\psi(\omega)$ is practically independent of ω and the imaginary part $Im\psi(\omega)$ is nearly zero. The experimental results justify the following mathematical assumption.

H5. For a dielectric, $Re\, \epsilon(\omega) = \epsilon_o + Re\psi(\omega)$ is independent of ω and the imaginary part $Im\psi(\omega)$ is zero, in a range $\omega_4 < \omega < \omega_5$ which is comprised between the infrared and ultraviolet absorption bands and corresponds to an interval

$$\lambda_o < \lambda < \lambda_1 \tag{6.41}$$

of wavelengths in vacuo $\lambda = 2\pi c_o/\omega$, with λ_o given by eq. (6.40).

We consider next the case of a conductor, such as a metal. The simple electron theory of metals envisages free electrons wandering among fixed ions

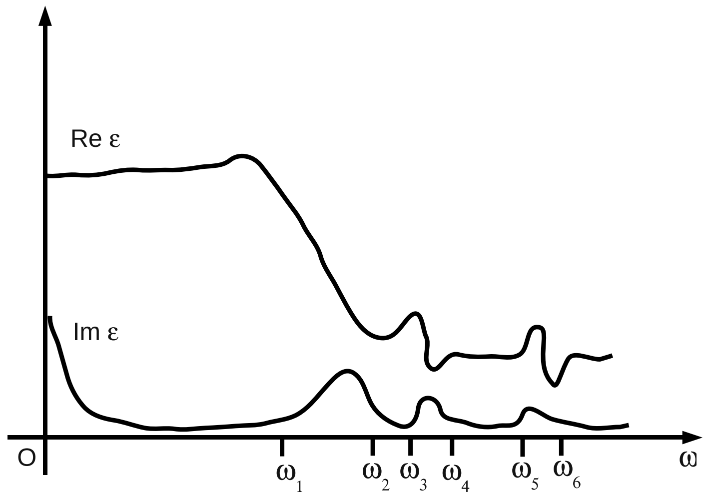

Figure 6.1: Behavior of $Re\,\epsilon(\omega)$ and of $Im\,\epsilon(\omega)$

and carrying the current, without any (average) restoring force (cfr. [42] p. 111). The equation of motion is the same as before, except that ω_o is replaced by zero:

$$\frac{d^2\mathbf{x}}{dt^2} + g\frac{d\mathbf{x}}{dt} = -\frac{e}{m}\boldsymbol{E}_o e^{i\omega t}$$

For a solution $\boldsymbol{x}$ of the form (6.34) we have

$$\omega\,(\omega - ig)\,\boldsymbol{x}_o = \frac{e}{m}\boldsymbol{E}_o$$

and the velocity $\boldsymbol{v} := d\boldsymbol{x}/dt$ is equal to $\boldsymbol{v}_o e^{i\omega t}$, where $\boldsymbol{v}_o = i\omega\boldsymbol{x}_o$ satisfies the equation

$$i\omega\boldsymbol{v}_o + g\boldsymbol{v}_o = -\frac{e}{m}\boldsymbol{E}_o$$

so that

$$\boldsymbol{v}_o = -\frac{e}{m}\frac{1}{g + i\omega}\boldsymbol{E}_o$$

The current density at the frequency ω will then be given by

$$\boldsymbol{J} = -N'e\boldsymbol{v} \equiv -\frac{N'e^2}{m}\frac{1}{g + i\omega}\boldsymbol{E}_o e^{i\omega t}$$

where $N' \leq N$ is the number of conduction electrons per unit volume. Since $\boldsymbol{J}$ is also equal to $\gamma(\omega)\boldsymbol{E}_o e^{i\omega t}$, we find

$$\gamma(\omega) = \frac{N'e^2}{m}\frac{1}{g + i\omega} \tag{6.42}$$

Thus $\gamma(\omega)$ has a simple pole at $\omega = ig$, and [3]

$$Re\gamma(\omega) = \frac{N'e^2}{m}\frac{g}{g^2+\omega^2} > 0 \quad , \quad Im\gamma(\omega) = -\frac{N'e^2}{m}\frac{\omega}{g^2+\omega^2} < 0 \tag{6.43}$$

Assumptions **H2**, **H3** and **H4** are satisfied, and the static conductivity is

$$\gamma(0) = \frac{N'e^2}{mg}$$

so that

$$g = N'e^2/m\gamma(0)$$

For a good conductor we may take $N' = N$, whereas for a bad conductor $N' \ll N$. For a common metal $\gamma(0)$ is of the order of $10^7 mho/m$ [22] and taking $N' = N$ we obtain the estimate $g \cong 10^{14} sec^{-1}$. Eq. (6.42) implies that the conductivity memory function is exponential

$$\tilde{\gamma}(t) = \sqrt{2\pi}\,\frac{N'e^2}{m}\,e^{-gt} \;\; \text{for } t > 0\,; \quad \tilde{\gamma}(t) \equiv 0 \;\; \text{for } t < 0$$

(Exercise 4).

In this simple conductor model, we may assume that the complex polarizability at the frequency ω is given by eq. (6.36) with N replaced by $N - N'$

$$\psi(\omega) = -\frac{e^2}{m}\frac{N - N'}{\omega^2 - \omega_o^2 - i\omega g}$$

so that the polarization memory function is given by eqs. (6.38) or (6.39) with N replaced by $N - N'$, and $\psi(\omega) \cong 0$ for $N' \cong N$ (good conductor), $\gamma(\omega) \cong 0$ for $N' \ll N$ (bad conductor, cfr. H3).

Remark 4. The constitutive relation (6.6) can be written in the form

$$\boldsymbol{J}(\boldsymbol{x},\, t) = \gamma_o \boldsymbol{E}(\boldsymbol{x},\, t) + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \gamma'(\tau)\boldsymbol{E}(\boldsymbol{x},\, t-\tau)d\tau$$

[3] $Re\tilde{\gamma}(\omega)$ is the (real) conductivity and $Im\tilde{\gamma}(\omega)$ the susceptivity

where $\gamma_o = \gamma(0)$ and $\gamma'(\tau)$ is the inverse Fourier transform of $\gamma(\omega) - \gamma(0)$, where

$$\gamma(\omega) - \gamma(0) \ = -\frac{N'e^2(\omega^2 + ig\omega)}{mg\,(\omega^2 + g^2)}$$

6.3 Approximation by local constitutive relations.

Time-harmonic fields such as (6.9), (6.11) and (6.13) are obviously related to Fourier trasforms. Consider the Fourier transforms of the electric field $\boldsymbol{E}$ and of the displacement vector $\boldsymbol{D}$

(6.44)

$$\tilde{\boldsymbol{E}}(\boldsymbol{x},\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-i\omega t}\boldsymbol{E}(\boldsymbol{x},t)dt\,,\ \ \tilde{\boldsymbol{D}}(\boldsymbol{x},\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-i\omega t}\boldsymbol{D}(\boldsymbol{x},t)dt$$

and the inversion formulae

(6.45)

$$\boldsymbol{E}(\boldsymbol{x},t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega t}\tilde{\boldsymbol{E}}(\boldsymbol{x},\omega)d\omega\ ,\ \ \boldsymbol{D}(\boldsymbol{x},t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega t}\tilde{\boldsymbol{D}}(\boldsymbol{x},\omega)d\omega$$

Since $\boldsymbol{D}$ is related to $\boldsymbol{E}$ by the hereditary relation (6.4)

$$\boldsymbol{D}(\boldsymbol{x},t) = \ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\tilde{\epsilon}(\tau)\boldsymbol{E}(\boldsymbol{x},\ t-\tau)d\tau$$

setting $\eta := t - \tau$ we find (convolution theorem)

$$(6.46)\qquad \tilde{\boldsymbol{D}}(\boldsymbol{x},\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\omega t}dt\int_{-\infty}^{\infty}\tilde{\epsilon}(\tau)\boldsymbol{E}(\boldsymbol{x},\ t-\tau)d\tau$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\omega\tau}\tilde{\epsilon}(\tau)d\tau\int_{-\infty}^{\infty} e^{-i\omega\eta}\boldsymbol{E}(\boldsymbol{x},\ \eta)d\eta\ = \epsilon(\omega)\tilde{\boldsymbol{E}}(\boldsymbol{x},\omega)$$

where, as we have seen, $\epsilon(\omega)$ is the Fourier transform of the memory function $\tilde{\epsilon}(\tau)$. We suppose that all these Fourier integrals exist in some sense and that all the above formal passages are justified.

(Note that for convenience we are denoting the Fourier transforms of field vectors with a tilde, $\tilde{\boldsymbol{E}}(\boldsymbol{x},\omega), \tilde{\boldsymbol{D}}(\boldsymbol{x},\omega), \boldsymbol{...}$, and the Fourier transforms of constitutive functions without a tilde, $\epsilon(\omega), \gamma(\omega)$,...)

Similarly, the Fourier transform

$$\tilde{\boldsymbol{J}}(\boldsymbol{x},\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\omega t} \boldsymbol{J}(\boldsymbol{x},t)dt$$

of the current density

$$\boldsymbol{J}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\gamma}(\tau) \boldsymbol{E}(\boldsymbol{x},\, t-\tau)d\tau \tag{6.47}$$

satisfies

$$\tilde{\boldsymbol{J}}(\boldsymbol{x},\omega) = \gamma(\omega)\, \tilde{\boldsymbol{E}}(\boldsymbol{x},\omega) \tag{6.48}$$

The following result, independent of assumption **H5**, shows that local constitutive relations are acceptable for narrow frequency intervals inside regions of normal dispersion.

Proposition 6.3.1 *Suppose* $\tilde{\boldsymbol{E}}(\boldsymbol{x},\omega)$ *vanishes identically outside a small interval* $\overline{\omega} - h < \omega < \overline{\omega} + h$ *, where the dispersion is normal. Then the approximate local constitutive relations*

$$\boldsymbol{D}(\boldsymbol{x},t) = \epsilon \boldsymbol{E}(\boldsymbol{x},t), \qquad \boldsymbol{J}(\boldsymbol{x},\, t) = \gamma \boldsymbol{E}(\boldsymbol{x},\, t) \tag{6.49}$$

hold with

$$\epsilon = \epsilon(\overline{\omega})\,,\ \gamma = Re\gamma(\overline{\omega}) \quad , \qquad Im\epsilon(\overline{\omega}) \cong Im\gamma(\overline{\omega}) \cong 0$$

Proof. If h is small we have

$$Re\psi(\omega) \cong Re\psi(\overline{\omega})\ ,\ \ Im\psi(\omega) \cong 0 \quad \text{for } \overline{\omega} - h < \omega < \overline{\omega} + h$$

(see Fig. 6.1), so that

$$\epsilon(\omega) \cong \epsilon(\overline{\omega}) = \epsilon_o + Re\psi(\overline{\omega}) \tag{6.50}$$

and $\epsilon(\overline{\omega})$ is (approximately) real and positive. Eqs. (6.44)−(6.46) imply then that

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{x},t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \epsilon(\omega) \tilde{\boldsymbol{E}}(\boldsymbol{x},\omega)d\omega \\ &\cong\ \epsilon(\overline{\omega}) \frac{1}{\sqrt{2\pi}} \int_{\overline{\omega}-h}^{\overline{\omega}+h} e^{i\omega t} \tilde{\boldsymbol{E}}(\boldsymbol{x},\omega)\, d\omega = \epsilon \boldsymbol{E}(\boldsymbol{x},t) \end{aligned}$$

so that the corresponding memory function is given by eq. (6.26)

$$\tilde{\epsilon}(\tau) \cong \sqrt{2\pi}\,\epsilon\,\delta(\tau)$$

Similarly, from eqs. (6.47) and (6.48) we find

$$\boldsymbol{J}(\boldsymbol{x},\, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t}\, \tilde{\boldsymbol{J}}(\boldsymbol{x}, \omega) d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \gamma(\omega)\, \tilde{\boldsymbol{E}}(\boldsymbol{x}, \omega) d\omega$$
$$\cong \gamma(\overline{\omega}) \frac{1}{\sqrt{2\pi}} \int_{\overline{\omega}-h}^{\overline{\omega}+h} e^{i\omega t} \tilde{\boldsymbol{E}}(\boldsymbol{x}, \omega) d\omega = \gamma \boldsymbol{E}(\boldsymbol{x}, t)$$

where $\gamma = \gamma(\overline{\omega})$ is approximately real and positive.

In particular, for a monochromatic field (6.9), with $\omega = \omega_o$

$$\boldsymbol{E}(\boldsymbol{x}, t) = \mathbb{E}(\boldsymbol{x}) e^{i\omega_o t}$$

the Fourier transform of $\boldsymbol{E}$ is given by the Dirac distribution

$$\tilde{\boldsymbol{E}}(\boldsymbol{x}, \omega) = \sqrt{2\pi} \mathbb{E}(\boldsymbol{x}) \delta(\omega - \omega_o) \tag{6.51}$$

Thus

$$\tilde{\boldsymbol{D}}(\boldsymbol{x}, \omega) = \sqrt{2\pi} \epsilon(\omega) \mathbb{E}(\boldsymbol{x}) \delta(\omega - \omega_o) \equiv \sqrt{2\pi} \epsilon(\omega_o) \mathbb{E}(\boldsymbol{x}) \delta(\omega - \omega_o)$$

and $\boldsymbol{D}$ is given by eq. (6.11), with $\omega = \omega_o$. In particular for $\omega_o = 0$ the field is static

$$\boldsymbol{E} = \mathbb{E}(\boldsymbol{x}) \quad , \quad \boldsymbol{D} = \epsilon(0) \mathbb{E}(\boldsymbol{x})$$

with $\epsilon(0)$ given by eq. (6.17). In the case of a uniform (white) frequency distribution the fields turns out to be approximately monochromatic for $|t| \ll h^{-1}$ (Exercise 5 and Exercise 6).

H5 and Propositon 6.3.1 imply the following

Corollary 6.3.2 *In a dielectric the local constitutive relation (6.49)*

$$\boldsymbol{D}(\boldsymbol{x}, t) = \epsilon \boldsymbol{E}(\boldsymbol{x}, t)$$

holds with

$$\epsilon := \epsilon(\overline{\omega}),\ Im \epsilon(\overline{\omega}) \cong 0 \tag{6.52}$$

in the frequency range $\overline{\omega} - h < \omega < \overline{\omega} + h$ *corresponding to the interval of wavelengths in vacuo (6.41).*

Remark 5. For dielectrics, absorption can be neglected whenever this non-dispersive limit case of local constitutive relations applies.

6.4 Anisotropic dielectric media. Kerr effect

The dispersive constitutive relation (6.4) or, equivalently, eq. (6.11) holds only in isotropic media. For anisotropic dielectrics (crystals) the directions of the vectors $\boldsymbol{E}$ and $\boldsymbol{D}$ in general do not coincide and the permittivity is a complex-valued rank-2 tensor $\epsilon_{ij}(\omega)$ $(i, j = 1, 2, 3)$. Therefore eq. (6.46) must be replaced by

$$\tilde{D}_i(\boldsymbol{x}, \omega) = \sum_{j=1}^{3} \epsilon_{ij}(\omega)\tilde{E}_j(\boldsymbol{x}, \omega) \qquad (i = 1, 2, 3)$$

where $\tilde{D}_i$, $\tilde{E}_j$ are the cartesian components of the Fourier transforms of the displacement and electric field vectors [4]. Accordingly, the memory function for a crystal, i.e. the inverse Fourier transform of $\epsilon_{ij}(\omega)$, is a real-valued rank-2 tensor $\tilde{\epsilon}_{ij}(\tau)$, and by taking the inverse Fourier transform of the previous equation and proceeding as in eqs. (6.45) and (6.46) we find the nonlocal constitutive relation

$$D_i(\boldsymbol{x}, t) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \sum_{j=1}^{3} \tilde{\epsilon}_{ij}(\tau) E_j(\boldsymbol{x}, t - \tau) d\tau \qquad (i = 1, 2, 3) \tag{6.53}$$

which replaces eq. (6.4). Thermodynamic considerations show [9,20,35] that, for each value of τ, $\tilde{\epsilon}_{ij}(\tau)$ is a real symmetric tensor

$$\tilde{\epsilon}_{ij}(\tau) = \tilde{\epsilon}_{ji}(\tau) \tag{6.54}$$

and hence can be diagonalized by a suitable choice of coordinate axes, the principal axes of the tensor. It follows that the complex permittivity tensor is also symmetric

$$\epsilon_{ij}(\omega) = \epsilon_{ji}(\omega)$$

and is diagonalied by the same choice of axes. The three eigenvalues of the matrix $[\epsilon_{ij}(\omega)]$ are the principal values $\epsilon_1(\omega), \epsilon_2(\omega), \epsilon_3(\omega)$ of the complex permittivity at the frequency ω. Crystals can be classified, from the point of view of their optical properties, on the grounds of the number of distinct eigenvalues.

[4] piroelectric effects [35]will be neglected throughout

(i) If the eigenvalues coincide for all ω

$$\epsilon_1(\omega) = \epsilon_2(\omega) = \epsilon_3(\omega) := \epsilon(\omega)$$

the principal axes can be chosen arbitrarily, $\epsilon_{ij}(\omega) = \epsilon(\omega)\delta_{ij}$,the crystal behaves like a isotropic medium and is called cubic. The velocity of light inside the crystal is the same in all directions.

(ii) If two eigenvalues coincide, say

$$\epsilon_1(\omega) = \epsilon_2(\omega) := \epsilon_\perp(\omega) \quad , \quad \epsilon_3(\omega) := \epsilon_\parallel(\omega) \neq \epsilon_\perp(\omega)$$

the first two principal axes can be chosen arbitrarily in a plane orthogonal to the third and the crystal is called uniaxial. Such crystal are axially symmetric around the third principal axis, which is the sole preferred direction of the crystal and is called the optic axis.

(iii) Finally if all the three eigenvalues are different the crystal is called biaxial.

In cases (ii) and (iii) the speed of light depends on the propagation direction.

We will be concerned from now on with uniaxial crystals. In a uniaxial crystal two types of plane (homogeneous) waves are possible, an ordinary wave and an extraordinary wave. If the optic axis coincides with the x_3-axis with unit vector $\mathbf{c}_3$, the crystal behaves with respect to the ordinary wave as an isotropic medium with refractive index

$$n_{or} = \sqrt{\epsilon_\perp(\omega)/\epsilon_o}$$

so that the wavenumber of the ordinary wave is

$$\boldsymbol{p}_{or} = \frac{\omega}{c_\perp}\,\mathrm{k}, \quad c_\perp = \frac{c_o}{n_{or}}$$

for any propagation direction $\boldsymbol{k}$ ($|\boldsymbol{k}| = 1,$ see eqs. (4.9) and (4.16)). Let

$$\phi := cos^{-1}(\boldsymbol{k} \cdot \mathbf{c}_3)$$

denote the angle between $\boldsymbol{k}$ and the optic axis, and suppose to begin with that $\boldsymbol{k}$ and $\mathbf{c}_3$ are not parallel ($sin\phi \neq 0$). The plane containing $\boldsymbol{k}$ and $\mathbf{c}_3$ is called the principal plane.

For the ordinary wave, the index of refraction n_{or} is independent of ϕ. In contrast, for any given $\boldsymbol{k}$ and angle $\phi := cos^{-1}(\boldsymbol{k}\cdot\mathbf{c}_3)$, the wavenumber of the extraordinary wave is given by

$$\boldsymbol{p}_{ex} = p_{ex}\boldsymbol{n} \quad , \quad p_{ex} := \frac{\omega\, n_{ex}}{c_o} \tag{6.55}$$

where

$$n_{ex} := \Big[\frac{\epsilon_o sin^2\phi'}{\epsilon_{\|}(\omega)} + \frac{\epsilon_o cos^2\phi'}{\epsilon_{\perp}(\omega)}\Big]^{-\frac{1}{2}}$$

and $\boldsymbol{n}$ is the unit vector in the principal plane which makes an angle

$$\phi' := arctan(\frac{\epsilon_{\perp}(\omega)}{\epsilon_{\|}(\omega)} tan\phi)$$

with the optic axis. Thus both the index of refraction and the direction of propagation $\boldsymbol{n}$ of the extraordinary wave depend on ϕ.

The numbers n_{or} and n_{ex} are called the principal indices of refraction of the crystal. If $\boldsymbol{k}$ is orthogonal to the optic axis ($cos\,\phi = 0$) then

$$n_{ex} = \sqrt{\epsilon_{\|}(\omega)/\epsilon_o} \quad , \; cos\phi' = 0 \;, \; p_{ex} = \frac{\omega}{c_{\|}} \quad , \quad \boldsymbol{n} = \boldsymbol{k}$$

where $c_{\|} = c_o/n_{ex}$, so that the directions of the ordinary and extraordinary waves coincide. The same happens if $\boldsymbol{k}$ is parallel to the optic axis ($sin\phi = 0$), since then

$$n_{ex} = \sqrt{\epsilon_{\perp}(\omega)/\epsilon_o} \quad , \quad sin\phi' = 0 \,, \; p_{ex} = \frac{\omega}{c_{\perp}} \;, \; \boldsymbol{n} = \boldsymbol{k}$$

In all other cases ($0 < \phi < \pi/2$) the propagation directions of the ordinary and extraordinary waves are different ($\boldsymbol{n}\neq\boldsymbol{k}$). The case of waves propagating along the optic axis is an exception in the sense that $\phi = \phi' = 0$ implies $n_{ex} = n_{or}$ and $\boldsymbol{p}_{ex} = \boldsymbol{p}_{or}$, so that the ordinary and extraordinary waves coincide and their superposition yields in general an elliptically polarized wave.

As regards the polarization, both waves are (linearly) polarized, with $\boldsymbol{E}$ and $\boldsymbol{D}$ orthogonal or parallel to the principal plane for the ordinary or extraordinary wave, respectively.

Summarizing, for a uniaxial crystal both the ordinary and extraordinary waves can propagate along a principal axis and they coincide if, and only if, they propagate along the optic axis. In all other cases their propagation directions are different:in refraction problems, this gives rise to the phenomenon of birefringence. If a plane wave impinges on a crystal surface which is neither parallel nor perpendicular to the optic axis, two refracted waves arise in the crystal, an ordinary and an extraordinary one. While the wavenumber of the ordinary refracted wave lies in the plane of incidence, as for isotropic media (see §4.8), this is not true in general for the extraordinary refracted wave.

An isotropic body may become anisotropic under the influence of an electric field. This phenomenon goes under the name of Kerr effect. Suppose $\boldsymbol{E}= E_o\mathbf{c}_3$ is a strong constant field: we have then

$$\epsilon_{ij}(\omega) = \epsilon(\omega)\delta_{ij} + \alpha E_i E_j \equiv \epsilon(\omega)\delta_{ij} + \alpha E_o^2\,\delta_{i3}\delta_{j3} \quad (i, j = 1, 2, 3)$$

where α is a (small) real constant, so that $\epsilon_\perp(\omega) = \epsilon(\omega) + \alpha E_o^2\,, \epsilon_\parallel(\omega) = \epsilon(\omega)$, and the principal axes of the tensor ϵ_{ij} are given by $\mathbf{c}_3$ (the optic axis) together with any pair of orthogonal axes in the $(\mathbf{c_1},\mathbf{c}_2)-$plane. Thus the isotropic body behaves like a uniaxial crystal under the influence of the electric field, and the permittivity tensor depends nonlinearly on the field strength E_o. Similarly, a uniaxial crystal may behave like a biaxial one under the influence of a strong electric field [35].

If dispersion can be neglected, the quantities $\epsilon_{ij}(\omega)$, $\epsilon(\omega)$, $\epsilon_\perp(\omega)$, $\epsilon_\parallel(\omega)$ are real and independent of ω, and ω disappears in all above considerations. We have then

$$\epsilon_{ij}(\omega) = \epsilon_{ij} \quad , \quad \epsilon_{ij} \in \mathbb{R} \tag{6.56}$$

so that the constitutive relation (6.53) becomes

$$D_i(\boldsymbol{x},\, t) = \sum_{j=1}^{3} \epsilon_{ij} E_j(\boldsymbol{x},\, t) \quad (i = 1, 2, 3) \tag{6.57}$$

In vector notation, eq. (6.57) can be written in the form

$$\boldsymbol{D} = \hat{\boldsymbol{\epsilon}}\boldsymbol{E}$$

where $\hat{\boldsymbol{\epsilon}}$ is the real symmetric rank-two tensor with components ϵ_{ij}. In particular, **H5**, Proposition 6.3.1 and Corollary 6.3.2 can be extended to anisotropic dielectrics, and we have

Corollary 6.4.1 *The local constitutive relation (6.57), with*

$$\epsilon_{ij} = \epsilon_{ij}(\overline{\omega}) \, , \, Im\epsilon_{ij} = 0 \tag{6.58}$$

holds approximately for an anisotropic dielectric in the frequency range $\overline{\omega} - h < \omega < \overline{\omega} + h$ *corresponding to the wavelength interval (6.41)*

$$\lambda_o < \lambda < \lambda_1$$

for the wavelength in vacuo $\lambda = 2\pi c_o/\omega$. *Thus in this frequency range all components of the memory tensor are Dirac functions*

$$\tilde{\epsilon}_{ij}(\tau) = \sqrt{2\pi}\, \epsilon_{ij}\, \delta(\tau)$$

$(i, j = 1, 2, 3)$.

6.5 Plane waves in uniaxial crystals.

Consider a plane linearly polarized monochromatic wave

$$\boldsymbol{E} = \boldsymbol{E}_o e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})} \quad , \quad \boldsymbol{H} = \boldsymbol{H}_o e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})} \tag{6.59}$$

propagating in a homogeneous uniaxial crystal, having permittivity tensor $\hat{\boldsymbol{\epsilon}}$ and magnetic permeability $\mu = \mu_o$. We assume that the frequency $\omega > 0$ is given and that the amplitudes $\boldsymbol{E}_o$, $\boldsymbol{H}_o$ and the wavenumber $\boldsymbol{p}$ are real vectors. We recall that the wavelength in the medium is defined by

$$\lambda := \frac{2\pi}{|\boldsymbol{p}|} \tag{6.60}$$

By force of eq. (6.58), the harmonic Maxwell equations (4.11) (with $\gamma = 0$) become here

$$curl\, \boldsymbol{E} = -i\omega\mu \boldsymbol{H} \ , \ curl \boldsymbol{H} = i\omega \boldsymbol{D} \equiv i\omega\hat{\boldsymbol{\epsilon}}\boldsymbol{E}$$

where the permittivity tensor $\hat{\boldsymbol{\epsilon}}$ has components ϵ_{ij} , depending in general upon ω. Proceeding as in Chapter 4 we find the relations

$$\boldsymbol{H}_o = \frac{1}{\omega\mu}\, \boldsymbol{p} \wedge \boldsymbol{E}_o \ , \quad \frac{1}{\omega}\boldsymbol{H}_o \wedge \boldsymbol{p} = \boldsymbol{D}_o := \hat{\boldsymbol{\epsilon}}\boldsymbol{E}_o \tag{6.61}$$

which imply that $\boldsymbol{p}\cdot\boldsymbol{D}_o = \boldsymbol{p}\cdot\boldsymbol{H}_o = 0$. Thus the three vectors $\boldsymbol{p}, \boldsymbol{D}_o, \boldsymbol{H}_o$ are mutually orthogonal, and, since $\boldsymbol{H}_o$ is also perpendicular to $\boldsymbol{E}_o$, the vectors $\boldsymbol{p}, \boldsymbol{D}_o, \boldsymbol{E}_o$ and $\boldsymbol{S}_o = \boldsymbol{E}_o \wedge \boldsymbol{H}_o$ lie in a common plane orthogonal to $\boldsymbol{H}_o$.

Note that here the vector $\boldsymbol{E}_o$ is not always parallel to $\boldsymbol{D}_o$ and is not necessarily transversal, i.e. orthogonal to $\boldsymbol{p}$.

It is useful to consider the energy velocity for a plane wave propagation in an anisotropic medium. This concept has been introduced in §4.7.2. In an isotropic medium the energy velocity coincides with the phase velocity of the wave.

Definition 6.5.1 *The energy velocity of the wave is defined as the normalized Poynting vector*

$$\boldsymbol{v}_e = \frac{\boldsymbol{E} \wedge \boldsymbol{H}}{\boldsymbol{E} \cdot \boldsymbol{D}}$$

For the plane wave (6.59) it is easy to see that $\boldsymbol{v}_e$ is time-independent and can be written as

$$\boldsymbol{v}_e = v_e \boldsymbol{n} \ , \ v_e = \left| \frac{\boldsymbol{E}_o \wedge \boldsymbol{H}_o}{\boldsymbol{E}_o \cdot \hat{\boldsymbol{\epsilon}} \boldsymbol{E}_o} \right|$$

where the ray vector $\boldsymbol{n}$ coincides with the unit normal to the wavefronts in geometric optics (see Chapter 4).

Let us choose for convenience the principal axes of the crystal (x, y, z) as cartesian axes, with $\mathbf{c}_3$ the optic axis. We then have

$$\boldsymbol{E}_o = E_1\,\mathbf{c}_1 + E_2\,\mathbf{c}_2 + E_3\,\mathbf{c}_3 \quad , \quad \boldsymbol{H}_o = H_1\mathbf{c}_1 + H_2\mathbf{c}_2 + H_3\mathbf{c}_3$$

and

$$\boldsymbol{D}_o = \epsilon_\perp(E_1\mathbf{c}_1 + E_2\mathbf{c}_2) + \epsilon_\parallel E_3\,\mathbf{c}_3$$

Substituting the first equation (6.61) into the second we obtain

$$\hat{\boldsymbol{\epsilon}}\boldsymbol{E}_o = \frac{1}{\omega^2\mu}(p^2\boldsymbol{E}_o - \boldsymbol{p}\boldsymbol{p}\cdot\boldsymbol{E}_o)\,, \ \ p\cdot\hat{\boldsymbol{\epsilon}}\boldsymbol{E}_o = 0$$

These relations can be written as a linear homogeneous algebraic system in the unknown $\boldsymbol{E}_o$, whose determinant must be set equal to zero. This yields the Fresnel equation (Exercise 7)

$$(p^2 - \omega^2\epsilon_\parallel\mu)(p^2 - \omega^2\epsilon_\perp\mu)(\frac{\epsilon_\perp}{\epsilon_\parallel}(p_1^2 + p_2^2) + p_3^2 - \omega^2\epsilon_\perp\mu) = 0$$

$(p^2 = p_1^2 + p_2^2 + p_3^2)$. The (positive) roots of this equation are

$$p = \omega\sqrt{\epsilon_\parallel \mu}\ ,\ p = \omega\sqrt{\epsilon_\perp \mu}\ \ ,\ \boldsymbol{p} = p_{ex}\boldsymbol{n}$$

where p_{ex} and $\boldsymbol{n}$ are given by eq. (6.55). We are interested in the cases of propagation parallel and orthogonal to the optic axis (see §6.4), corresponding to the first two roots.

A. Propagation orthogonal to the optic axis. We may assume without loss of generality $\boldsymbol{p}= p\mathbf{c}_2$ and

$$\boldsymbol{H}_o = H_1\mathbf{c}_1 + H_3\mathbf{c}_3\ ,\quad \boldsymbol{D}_o = D_1\mathbf{c}_1 + D_3\mathbf{c}_3\ ,\ \boldsymbol{E}_o = \epsilon_\perp^{-1}D_1\mathbf{c}_1 + \epsilon_\parallel^{-1}D_3\mathbf{c}_3$$

Eqs. (6.61) yield then

$$\omega D_1 = -pH_3\ ,\ \ \omega D_3 = pH_1\ ,\ \omega\mu H_1 = p\epsilon_\parallel^{-1}D_3\ ,\ \omega\mu H_3 = -p\epsilon_\perp^{-1}D_1$$

Since the $(x, y)-$axes can be arbitrarily rotated, two types of waves, with different phase velocities, are possible:

(i) Ordinary wave: $p = \omega\sqrt{\epsilon_\perp\mu}$, $H_3 = -D_1/\sqrt{\epsilon_\perp\mu}$ and

$$\boldsymbol{H}_o = H_3\mathbf{c}_3 \equiv -\frac{D_1}{\sqrt{\epsilon_\perp\mu}}\mathbf{c}_3\ ,\ \boldsymbol{D}_o = D_1\mathbf{c}_1\ ,\ \ \boldsymbol{E}_o = \epsilon_\perp^{-1}D_1\mathbf{c}_1$$

This wave is polarized with $\boldsymbol{D}$ parallel to $\boldsymbol{E}$ and orthogonal to the optic axis, and has phase speed, refraction index and wavenumber given by

$$c_{or} := \frac{1}{\sqrt{\epsilon_\perp\,\mu}}\,,\ n_{or} = \sqrt{\frac{\epsilon_\perp}{\epsilon_o}}\ ,\ \boldsymbol{p} = \omega\sqrt{\epsilon_\perp\,\mu}\mathbf{c}_2\ \equiv\ \frac{\omega}{c_{or}}\mathbf{c}_2 \tag{6.62}$$

This wave is transversal, its phase and energy velocity coincide ($\boldsymbol{v}_f = c_{or}\mathbf{c}_2 \equiv \boldsymbol{v}_e$), and the group velocity $\boldsymbol{v}_g = \omega'(p)\,\boldsymbol{k}$ is also parallel to $\mathbf{c}_2$.

(ii) Extraordinary wave: $p = \omega\sqrt{\epsilon_\parallel\mu}$, $H_1 = D_3/\sqrt{\epsilon_\parallel\mu}$ and

$$\boldsymbol{H}_o = H_1\mathbf{c}_1 \equiv \frac{D_3}{\sqrt{\epsilon_\parallel\mu}}\mathbf{c}_1\ ,\ \boldsymbol{D}_o = D_3\mathbf{c}_3\ ,\ \boldsymbol{E}_o = \epsilon_\parallel^{-1}D_3\mathbf{c}_3 \tag{6.63}$$

This wave is transversal, with $\boldsymbol{D}$ and $\boldsymbol{E}$ both parallel to the optic axis, and has phase speed, refraction index and wavenumber

$$c_{ex} := \frac{1}{\sqrt{\epsilon_\parallel\mu}}\ ,\ \ n_{ex} = \sqrt{\frac{\epsilon_\parallel}{\epsilon_o}}\ ,\ \ \boldsymbol{p} = \omega\sqrt{\epsilon_\parallel\mu}\,\boldsymbol{n}\ \equiv\ \frac{\omega}{c_{ex}}\mathbf{c}_2 \tag{6.64}$$

Thus $\boldsymbol{n}=\mathbf{c}_2$, the velocities $\boldsymbol{v}_f = c_{ex}\mathbf{c}_2 \equiv \boldsymbol{v}_e$, $\boldsymbol{v}_g = \omega'(p)\mathbf{c}_2$ are all parallel to $\mathbf{c}_2$, and the wavelength in the crystal is

$$\lambda_{ex} := \frac{2\pi c_o}{\omega n_{ex}}$$

In this case the ordinary and extraordinary waves propagate in the same direction but with different speeds.

B. Propagation parallel to the optic axis: $\boldsymbol{p}= p\mathbf{c}_3$. Here we may assume without loss of generality

$$\boldsymbol{H}_o = H_2\mathbf{c}_2 \quad \boldsymbol{D}_o = D_1\mathbf{c}_1 \ , \ \boldsymbol{E}_o = \epsilon_\perp^{-1} D_1\mathbf{c}_1$$

Eqs. (6.61) yield then

$$\omega\, D_1 = pH_2 \ , \ \omega\mu H_2 = p\,\epsilon_\perp^{-1} D_1$$

whence $p = \omega\sqrt{\epsilon_\perp\mu}$, $H_2 = D_1/\sqrt{\epsilon_\perp\mu}$ $(\mu = \mu_o)$ and

$$\boldsymbol{p} = \frac{\omega}{c_{or}}\,\mathbf{c}_3 \ \equiv \ \frac{\omega\, n_{or}}{c_o}\mathbf{c}_3$$

where

$$c_{or} := \frac{1}{\sqrt{\epsilon_\perp\mu}} \ , \ n_{or} = \sqrt{\frac{\epsilon_\perp}{\epsilon_o}} \ , \ \lambda_{or} := \frac{2\pi c_o}{\omega n_{or}} \qquad (6.65)$$

are the phase speed, index of refraction and length of the wave, respectively (Exercise 8). Since the $(x, y)-$axes can be arbitrarily rotated, this wave is transversal and arbitrarily polarized, with $\boldsymbol{D}$ parallel to $\boldsymbol{E}$, and $\boldsymbol{v}_e$ coincides with the phase velocity $\boldsymbol{v}_f = c_{or}\boldsymbol{k}$. Thus $\boldsymbol{n}=\mathbf{c}_3$ and the phase velocity, the energy velocity and the group velocity $\boldsymbol{v}_g = \omega'(p)\,\boldsymbol{k}$ are all parallel to the wavenumber $\boldsymbol{p}$.

In this case the ordinary and the extraordinary waves coincide.

6.6 Anisotropic conductors. Hall effect.

For anisotropic conductors the directions of the vectors $\boldsymbol{E}$ and $\boldsymbol{J}$ in general do not coincide and the conductivity is a complex-valued rank-2 tensor $\gamma_{ij}(\omega)$, so that eq. (6.13) must be replaced by

$$J_i = \sum_{j=1}^{3} \gamma_{ij}(\omega) E_j \qquad (i = 1, 2, 3) \tag{6.66}$$

The conductivity memory function, the inverse Fourier transform of $\gamma_{ij}(\omega)$, is a real-valued rank-2 tensor $\tilde{\gamma}_{ij}(\tau)$ and eq. (6.6) for an anisotropic conductor becomes

$$J_i(\boldsymbol{x}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_{j=1}^{3} \tilde{\gamma}_{ij}(\tau) E_j(\boldsymbol{x}, t - \tau) d\tau \tag{6.67}$$

Due to Onsager's symmetry principle [35], $\tilde{\gamma}_{ij}(\tau)$ is a real symmetric tensor

$$\tilde{\gamma}_{ij}(\tau) = \tilde{\gamma}_{ji}(\tau) \tag{6.68}$$

for each value of τ, hence the complex conductivity tensor is also symmetric

$$\gamma_{ij}(\omega) = \gamma_{ji}(\omega) \tag{6.69}$$

for every ω.

In the presence of a magnetic field $\boldsymbol{H}$ the conductivity tensor may become a function of $\boldsymbol{H}$ and in this case Onsager's principle implies that

$$\gamma_{ij}(\omega, \boldsymbol{H}) = \gamma_{ji}(\omega, -\boldsymbol{H})$$

hence $\gamma_{ij}(\omega,\boldsymbol{H})$ is no longer symmetric. It follows that $\boldsymbol{J}$ can be decomposed into a symmetric and an antisymmetric part in the form

$$\boldsymbol{J} = \mathbb{S}\boldsymbol{E} + \boldsymbol{E} \wedge \mathbf{a}(\omega, \boldsymbol{H}) \tag{6.70}$$

where $\mathbb{S}$ is the 3×3 matrix with entries $s_{ij}(\omega,\boldsymbol{H}) := \frac{1}{2}(\gamma_{ij}(\omega,\boldsymbol{H}) + \gamma_{ji}(\omega,\boldsymbol{H}))$, the symmetric part of the tensor $\gamma_{ij}(\omega,\boldsymbol{H})$, satisfying

$$s_{ji}(\omega, \boldsymbol{H}) = s_{ji}(\omega, -\boldsymbol{H}) = s_{ij}(\omega, -\boldsymbol{H}) \quad (i, j = 1, 2, 3)$$

and $\boldsymbol{a}(\omega,\boldsymbol{H})$ is a vector function with cartesian components $a_i(\omega,\boldsymbol{H})$ satisfying

$$a_i(\omega, \boldsymbol{H}) = -a_i(\omega, -\boldsymbol{H}) \quad (i = 1, 2, 3)$$

Thus $s_{ij}(\omega,\boldsymbol{H})$ and $a_i(\omega,\boldsymbol{H})$ are even and odd functions of $\boldsymbol{H}$, respectively so that

$$a_i(\omega,\boldsymbol{H}) \cong \sum_{k=1}^{3} \alpha_{ik}(\omega) H_k$$
$$s_{ij}(\omega,\boldsymbol{H}) \cong \gamma^o_{ij}(\omega)$$

up to terms of order $O(|\boldsymbol{H}|^2)$ as $|\boldsymbol{H}| \to 0$. The first relation can be written in vector form

$$\boldsymbol{a}(\omega,\boldsymbol{H}) \cong \mathbb{A}(\omega)\boldsymbol{H}$$

where $\mathbb{A}(\omega)$ is the 3×3 matrix with entries $\alpha_{ik}(\omega)$, and eq. (6.70) takes the form

$$\boldsymbol{J} = \mathbb{S}\boldsymbol{E} + \boldsymbol{E}\wedge \mathbb{A}(\omega)\boldsymbol{H} \tag{6.71}$$

Note that the Joule heat

$$\boldsymbol{E}\cdot\boldsymbol{J} = \sum_{i,j=1}^{3} s_{ij}(\omega,\boldsymbol{H})E_iE_j \cong \sum_{i,j=1}^{3} \gamma^o_{ij}(\omega)E_iE_j$$

depends solely on s_{ij} and is independent of $\boldsymbol{H}$. In practice the (moduli of the) coefficients $\alpha_{ik}(\omega)$ are small so that the Hall effect can often be neglected, and if not, the above relations can be assumed to be universally valid except for very strong magnetic fields.

Summarizing, eq. (6.71) shows that the main influence of the magnetic field consists in the appearance of a nonlinear electric current $\boldsymbol{E}\wedge\mathbb{A}(\omega)\boldsymbol{H}$ perpendicular to the electric field and linear in $\boldsymbol{H}$ (Hall effect). Conversely, one easily sees that

$$\boldsymbol{E} = \mathbb{S}^{-1}\boldsymbol{J} + \boldsymbol{J}\wedge \boldsymbol{b}(\omega,\boldsymbol{H}) \tag{6.72}$$

where $\boldsymbol{b}(\omega,\boldsymbol{H})$ is (approximately) a linear vector function of $\boldsymbol{H}$. Thus for given $\boldsymbol{J}$ the Hall effect consists in the appearance of a nonlinear electric field $\boldsymbol{J}\wedge\boldsymbol{b}(\omega,\boldsymbol{H})$ perpendicular to the current and depending linearly on $\boldsymbol{H}$.

Remark 6. If the conductor is isotropic the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ must be parallel to $\boldsymbol{H}$ and we have

$$\boldsymbol{E} = \frac{1}{\gamma(\omega)}\boldsymbol{J} + R(\omega)\boldsymbol{H}\wedge\boldsymbol{J}$$

where $R(\omega)$ is the Hall constant. Thus the Hall effect in an isotropic conductor gives rise to the appearance of an electric field $R(\omega)\boldsymbol{H}\wedge\boldsymbol{J}$ orthogonal to both the current and the magnetic field, and Ohm's law becomes nonlinear.

If dispersion can be neglected (see **H5** and Proposition 6.3.1 above), all the scalar, vector and tensor quantities $\gamma\,(\omega)$, $R(\omega)$, $\gamma_{ij}(\omega)$.... defined above can be considered real and independent of ω.

6.7 Second-harmonic generation in nonlinear optics.

Nonlinear properties of Maxwell's constitutive relations for the electric field $\boldsymbol{D}=\boldsymbol{D}(\boldsymbol{E})$ in a dielectric have been known from the beginning. The permittivity (tensor) is in general a function of the electric field $\boldsymbol{E}$, and the displacement vector $\boldsymbol{D}$ depends nonlinearly upon $\boldsymbol{E}$. For anisotropic dielectrics (crystals), this nonlinear dependence will be nonlocal. But, since the optical nonlinearities are small, their experimental discovery had to wait for the development of powerful lasers. Thus, it was only in 1961 that Franken, Ward and coworkers first detected ultraviolet light (with wavelength 3471Å) at twice the frequency of a ruby red laser beam (with wavelength 6943 Å), when this beam traversed a quartz crystal [23]. This experiment marked the beginning of an intensive research activity in both experimental and theoretical nonlinear optical properties.

6.7.1 The nonlinear constitutive relations.

The theory of wave propagation in nonlinear dielectrics can be developed along classical lines, using a model based on the Maxwell equations with nonlinear constitutive relations for the electric field [26]. Since the nonlinearity is weak, the first correction to the linear dependence of $\boldsymbol{D}$ on $\boldsymbol{E}$ is quadratic and for an anisotropic dielectric (crystal) it has the form [35]
(6.73)

$$D_i'(\boldsymbol{x},\,t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sum_{j,k=1}^{3}\tilde{\psi}_{ijk}(\tau_1,\tau_2)E_j(\boldsymbol{x},\,t-\tau_1)E_k(\boldsymbol{x},\,t-\tau_2)d\tau_1 d\tau_2$$

$(i = 1, 2, 3)$. Here $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ is a third-rank real tensor function of (τ_1, τ_2), vanishing identically for materials with a center of symmetry, as e.g. isotropic media, and different from zero for a crystal without inversion symmetry, i.e. a piezoelectric crystal [41]. For piezoelectric crystals the subsequent nonlinear terms in the relation $\boldsymbol{D}=\boldsymbol{D(E)}$ can be neglected, and the constitutive relation (6.53) must be replaced by

$$(6.74) \qquad D_i(\boldsymbol{x},\ t) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \sum_{j=1}^{3} \tilde{\epsilon}_{ij}(\tau) E_j(\boldsymbol{x},\ t-\tau) d\tau + \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{j,k=1}^{3} \tilde{\psi}_{ijk}(\tau_1, \tau_2) E_j(\boldsymbol{x},\ t-\tau_1) E_k(\boldsymbol{x},\ t-\tau_2) d\tau_1 d\tau_2$$

$(i = 1, 2, 3)$. The double integral in (6.74) represents the contribution due to the nonlinear polarization, and $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ is its memory (tensor) function. The causality principle implies that $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ satisfies

$$\tilde{\psi}_{ijk}(\tau_1, \tau_2) \equiv 0 \quad \text{for } \tau_1 < 0 \ \text{ or } \tau_2 < 0$$

so that the double integral in (6.74) goes from 0 to ∞. Moreover, interchanging τ_1 and τ_2 in eq. (6.74) yields the symmetry relation

$$(6.75) \qquad \tilde{\psi}_{ijk}(\tau_1, \tau_2) = \tilde{\psi}_{ikj}(\tau_1, \tau_2) \qquad \text{(for all } i, j, k, \tau_1, \tau_2)$$

Suppose that in eq. (6.74) $\boldsymbol{E}$ is given by the plane monochromatic wave (6.59)

$$\boldsymbol{E}(\boldsymbol{x},\ t) = Re\Big[\boldsymbol{E}_o e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})}\Big]$$

where $\boldsymbol{p}= p\boldsymbol{k}$ and $\boldsymbol{E}_o$ is a real vector (we recall that, since the relation (6.74) is nonlinear,we must take the real part of $\boldsymbol{E}$ in advance). An easy calculation yields

$$(6.76) \quad D_i(\boldsymbol{x},\ t) = Re\Big[\sum_{j=1}^{3} E_{oj}\epsilon_{ij}(\omega)\, e^{i(\omega t - \boldsymbol{p}\cdot\boldsymbol{x})} + \frac{1}{2} \sum_{j,k=1}^{3} E_{oj}E_{ok}\psi_{ijk}(\omega, -\omega) + \frac{1}{2} \sum_{j,k=1}^{3} E_{oj}E_{ok}\psi_{ijk}(\omega, \omega)\, e^{i(2\omega t - 2\boldsymbol{p}\cdot\boldsymbol{x})}\Big]$$

where E_{oj} are the cartesian components of $\boldsymbol{E}_o$, and the 2D Fourier transform of $\tilde{\psi}_{ikj}(\tau_1, \tau_2)$

$$\psi_{ijk}(\omega_1, \omega_2) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i(\omega_1\tau_1+\omega_2\tau_2)} \tilde{\psi}_{ijk}(\tau_1, \tau_2)\, d\tau_1 d\tau_2$$

is the nonlinear polarizability at the frequencies (ω_1, ω_2). We remark that eq. (6.75) and the fact that the memory function $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ is real imply the symmetry relations

$$\psi_{ijk}(\omega_1, \omega_2) = \psi_{ikj}(\omega_2, \omega_1)\ , \ \ \psi_{ijk}(-\omega_1, -\omega_2) = \psi^*_{ijk}(\omega_1, \omega_2) \tag{6.77}$$

(for all i, j, k, ω_1, ω_2),where the star * denotes the complex conjugate. (These relations imply that the second sum on the right-hand-side of eq. (6.76) is actually real.) Thus if $\boldsymbol{E}$ is a plane monochromatic wave in all space, the nonlinear polarization integral gives rise to a second-harmonic wave

$$Re\Big[\frac{1}{2} \sum_{j,k=1}^{3} E_{oj} E_{ok} \psi_{ijk}(\omega, \omega)\, e^{i(2\omega t - 2\boldsymbol{p}\cdot\boldsymbol{x})}\Big] \tag{6.78}$$

at twice the frequency ω of the original plane wave. Note that (6.78) is a traveling wave with phase velocity ω/p, which in empty space or air equals c_o, periodic in t with period π/ω (hence also with period $2\pi/\omega$).

This simple calculation, although far from representing an exact solution of the nonlinear Maxwell equations, shows how the quadratic term (6.73) gives rise to second-harmonic generation (abbreviated SHG). In fact, SHG is observed experimentally in piezoelectric crystals, like quartz. This is the class of nonlinear dielectrics to which we restrict our attention from now on.

The experimental results for the nonlinear polarizability of piezoelectric crystals [5] suggest the following ad hoc mathematical assumption.

H6. For a piezoelectric crystal $\psi_{ijk}(\omega_1, \omega_2)$ is real and independent of (ω_1, ω_2)

$$\psi_{ijk}(\omega_1, \omega_2) = \psi_{ijk}$$

when both frequencies ω_1, ω_2 are in the range $(\overline{\omega} - h, \overline{\omega} + h)$ corresponding to wavelengths in vacuo $\lambda = 2\pi c_o/\omega$ comprised in the interval (6.41).

[5] The experimental values reported in [24] show a slight dependence of ψ_{ijk} on frequency in the range between the infrared and ultraviolet absorption bands

According to **H5, H6** and the Corollaries above, all components of the memory tensors $\tilde{\epsilon}_{ij}(\tau)$, $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ are Dirac distributions

$$\tilde{\epsilon}_{ij}(\tau) = \sqrt{2\pi}\, \epsilon_{ij}\, \delta(\tau) \;\; , \;\; \tilde{\psi}_{ijk}(\tau_1, \tau_2) = 2\pi \psi_{ijk} \delta(\tau_1)\delta(\tau_2) \tag{6.79}$$

$(i, j, k = 1, 2, 3)$ in the wavelength range $\lambda_o < 2\pi c_o/\omega < \lambda_1$. It turns out that this is the appropriate wavelength range in the experiments concerning SHG. (For example, in the original experiments by Franken et al. the incident wave from the ruby laser beam has wavelength 6943 Å and the second-harmonic wave has wavelength 3471Å.) One can then assume the approximate nonlinear local constitutive relation between $\boldsymbol{D}$ and $\boldsymbol{E}$ in a (piezoelectric) crystal

$$D_i(\boldsymbol{x}, t) = \sum_{j=1}^{3} \epsilon_{ij} E_j(\boldsymbol{x}, t) + \sum_{j,k=1}^{3} \psi_{ijk} E_j(\boldsymbol{x}, t) E_k(\boldsymbol{x}, t) \tag{6.80}$$

with $\epsilon_{ij} := \epsilon_{ij}(\overline{\omega})$, $\psi_{ijk} := \psi_{ijk}(\overline{\omega}, \overline{\omega})$, $Im\, \epsilon_{ij} = Im \psi_{ijk} = 0$ $(i, j, k = 1, 2, 3)$. In vector form this equation becomes

$$\boldsymbol{D} = \hat{\boldsymbol{\epsilon}} \boldsymbol{E} + \hat{\boldsymbol{\psi}} \boldsymbol{E}\boldsymbol{E} \tag{6.81}$$

where $\hat{\boldsymbol{\epsilon}}$ is the real symmetric rank-two tensor having constant components ϵ_{ij} and $\hat{\boldsymbol{\psi}}$ is the real rank-3 tensor having constant components ψ_{ijk}. As concerns the magnetic field, the linear local relation

$$\boldsymbol{B} = \mu \boldsymbol{H}$$

can be retained with $\mu = \mu_o$. With the aid of these constitutive relations we arrive at the following nonlinear system of Maxwell equations (with $\boldsymbol{J} = \rho = \sigma = 0$, $\mu = \mu_o$ everywhere)

$$\mu_o \frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E}\, , \quad \frac{\partial}{\partial t}(\hat{\boldsymbol{\epsilon}} \boldsymbol{E} + \hat{\boldsymbol{\psi}} \boldsymbol{E}\boldsymbol{E}) = curl\, \boldsymbol{H} \tag{6.82}$$

$$div\, (\hat{\boldsymbol{\epsilon}} \boldsymbol{E} + \hat{\boldsymbol{\psi}} \boldsymbol{E}\boldsymbol{E}) = 0\, , \quad div \boldsymbol{H} = 0 \tag{6.83}$$

inside the crystal. Outside the crystal, in empty space (or air), the Maxwell equations are the usual linear ones

$$\mu_o \frac{\partial \boldsymbol{H}}{\partial t} = -curl\, \boldsymbol{E}\, , \qquad \epsilon_o \frac{\partial \mathbf{E}}{\partial t} = curl\, \boldsymbol{H} \tag{6.84}$$

$$div\,\boldsymbol{E} = 0 \quad , \quad div\,\boldsymbol{H} = 0 \tag{6.85}$$

Finally, the appropriate matching conditions must hold across the crystal surface $\mathbb{S}$. Since we are assuming $\mu = \mu_o$ and $\sigma = 0$ everywhere, the tangential components of $\boldsymbol{E}$, the normal component of $\boldsymbol{D}= \hat{\boldsymbol{\epsilon}}\boldsymbol{E}+\hat{\boldsymbol{\psi}}\boldsymbol{EE}$ and the magnetic field $\boldsymbol{H}$ must be continuous across $\mathbb{S}$ (see §1.4):

$$[\boldsymbol{E}]_{\mathbb{S}} \wedge \boldsymbol{n} = \boldsymbol{0} \quad , \quad [\boldsymbol{H}]_{\mathbb{S}} = \boldsymbol{0}, \quad [\boldsymbol{D}]_{\mathbb{S}} \cdot \boldsymbol{n} = 0 \tag{6.86}$$

The solutions of the Maxwell equations inside and outside the crystal must be matched via these relations, so that the problem of SHG can be formulated mathematically as a transmission problem for the nonlinear Maxwell equations. For simple geometries, such as those adopted in the experiments by Franken and Ward [23,24], the transmission problem is equivalent to a pure boundary value problem for the nonlinear Maxwell equations in one space variable, as in the case of the linear problem considered in §4.9.

6.7.2 The "laser problem".

We shall consider the following mathematical model for second-harmonic generation: A plane linearly polarized monochromatic harmonic wave travels in air (or vacuo) along the direction of a unit vector $\boldsymbol{k}$ in the region $\boldsymbol{k}\cdot\boldsymbol{x}< 0$

$$\boldsymbol{E} = \boldsymbol{E}_o\, cos\,(\omega t - \frac{\omega}{c_o}\boldsymbol{k}\cdot\boldsymbol{x}) \ , \quad \boldsymbol{H} = \boldsymbol{H}_o\, cos\,(\omega t - \frac{\omega}{c_o}\boldsymbol{k}\cdot\boldsymbol{x}) \tag{6.87}$$

and impinges "to the left" on a piezoelectric crystal slab D_a of (small) thickness a

$$D_a = \left\{0 \le \boldsymbol{k}\cdot\boldsymbol{x} \le a\right\}$$

at normal incidence. The continuous incident wave (6.87) is meant to approximate the beam emitted by a laser:this is acceptable even in the case of pulsed lasers, since the duration of a laser pulse is several orders of magnitude higher than its period.

We assume that the crystal is uniaxial and the slab is cut orthogonally to one of the principal axes, so that the phase velocity $\frac{\omega}{c_o}\boldsymbol{k}$ of the incident wave (6.87) is parallel to a principal axis of the crystal. When traversing the crystal slab, this incident wave undergoes a process of multiple reflections at the two plane slab interfaces ($\boldsymbol{k}\cdot\boldsymbol{x}= 0$, a) due to the matching relations

(6.86), as explained at the beginning of §4.9. This process gives rise to a reflected wave in the first region $\boldsymbol{k}\cdot\boldsymbol{x}< 0$ and a transmitted wave in the third region $\boldsymbol{k}\cdot\boldsymbol{x}> a$. The permittivities ϵ_1 and ϵ_3 outside the crystal, in the first and third regions respectively, are the same as the permittivity of air, i.e. $\epsilon_1 = \epsilon_3 = \epsilon_o$ and $\boldsymbol{D}= \epsilon_o\boldsymbol{E}$ outside D_a, for all t. The electromagnetic field depends only on one scalar space variable $\boldsymbol{k}\cdot\boldsymbol{x}$:

$$\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{k}\cdot\boldsymbol{x}, t)\ , \quad \boldsymbol{H} = \boldsymbol{H}(\boldsymbol{k}\cdot\boldsymbol{x}, t)$$

and is transversal outside the crystal, so that the longitudinal components of $\boldsymbol{E}$ and $\boldsymbol{H}$ are zero outside the slab:

$$\boldsymbol{D}\cdot\boldsymbol{k} \equiv \epsilon_o\boldsymbol{E}\cdot\boldsymbol{k} = 0 \quad , \ \boldsymbol{H}\cdot\boldsymbol{k} = 0 \qquad \text{(outside } D_a)$$

As the unit normal $\boldsymbol{n}$ to the slab surfaces is parallel to $\boldsymbol{k}$, the second and third matching relation (6.86) yield $\boldsymbol{D}\cdot\boldsymbol{k}=\boldsymbol{H}\cdot\boldsymbol{k}= 0$ on the inner slab interfaces for all t, and eqs. (6.83)

$$div\,\boldsymbol{D}(\boldsymbol{k}\cdot\boldsymbol{x}, t) = div\,\boldsymbol{H}(\boldsymbol{k}\cdot\boldsymbol{x}, t) = 0$$

imply that $\boldsymbol{D}\cdot\boldsymbol{k}=\boldsymbol{H}\cdot\boldsymbol{k}= 0$ everywhere inside the crystal, with $\boldsymbol{D}= \hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E}$ (eq. (6.81)). Therefore the magnetic field must be transversal everywhere

$$\text{(6.88)} \qquad \boldsymbol{H}\cdot\boldsymbol{k} = 0 \qquad \text{(for all } \boldsymbol{x} \text{ and } t)$$

The electric field must satisfy the relations

$$\text{(6.89)} \qquad \boldsymbol{D}\cdot\boldsymbol{k} \equiv \epsilon_o\boldsymbol{E}\cdot\boldsymbol{k} = 0 \ \ \text{(outside } D_a)$$

$$\boldsymbol{D}\cdot\boldsymbol{k} \equiv \hat{\boldsymbol{\epsilon}}\boldsymbol{E}\cdot\boldsymbol{k} + (\hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})\cdot\boldsymbol{k} = 0 \ \ \text{(in } D_a)$$

and is not necessarily transversal inside the crystal (cfr. §6.5).

The rank-2 tensor $\hat{\boldsymbol{\epsilon}}$ is diagonal with eigenvalues $\epsilon_\perp$ and $\epsilon_\parallel$, while the entries of the rank-3 tensor ψ_{ijk} for the various crystals can be found from the specialized literature (see e.g. [41]). According to **H5, H6** and Corollary 6.4.1, we can neglect dispersion and think of $\epsilon_\perp$, $\epsilon_\parallel$ and ψ_{ijk} as real constants, independent of frequency:this assumption is reasonable provided the thickness a of the slab is small. The symmetry relations (6.77) then reduce to

$$\psi_{ijk} = \psi_{ikj}$$

(for all i, j, k).

A crucial fact is that, irrespective of the nature of the nonlinear dielectric slab, **the form of the solution outside the slab is known and is represented by transversal traveling waves** with phase velocity c_o , exactly as in §4.9. Precisely, since the Maxwell equations (6.84), (6.85) outside the crystal are linear, homogeneous, with constant coefficients and one space variable $\boldsymbol{k}\cdot\boldsymbol{x}$, the general solution will be a superposition of a progressive (incident) wave and a regressive (reflected) wave in the first region $\boldsymbol{k}\cdot\boldsymbol{x}< 0$, and a progressive (transmitted) wave in the third region $\boldsymbol{k}\cdot\boldsymbol{x}> a$, where no reflected wave is expected to occur (see Chapter 4). This fact has several important consequences:

(i) The given incident traveling wave (6.87) emitted by the laser, must be transversal, i.e. $\boldsymbol{E}_o$ and $\boldsymbol{H}_o$ must be orthogonal to the unit vector $\boldsymbol{k}$. Indeed, if

(6.90) $$\boldsymbol{H}_o = \sqrt{\frac{\epsilon_o}{\mu_o}}\ \boldsymbol{k} \wedge \boldsymbol{E}_o \quad , \ \boldsymbol{k} \cdot \boldsymbol{E}_o = 0$$

the incident wave (6.87) is a periodic solution of the Maxwell equations (6.84), (6.85) (of period $2\pi/\omega$) for any propagation direction $\boldsymbol{k}$ (see §4.2).

(ii) The (unknown) reflected and transmitted waves must also be transversal traveling waves.

(iii) Although the reflected and transmitted waves are unknown, the known form of the solution outside the slab D_a combined with the matching relations (6.86) determines a set of **exact boundary conditions** *on the two slab boundaries for the electromagnetic field inside the slab.* As in the linear isotropic case (see Exercise 15 of Chapter 4), these boundary conditions, of impedance type, are linear and non-homogeneous, with the the incident wave as a driving term, and depend only on the form of the solution outside the crystal slab, irrespective of what takes place inside. (In particular they still hold even if dispersion in the crystal is taken into account and the nonlinear Maxwell equations (6.82), (6.83) in D_a are replaced by different equations, as e.g. integrodifferential ones.) In this way the problem of SHG can be reduced to solving a boundary value problem for a nonlinear hyperbolic system in characteristic form in the slab D_a, as first recognized by Cesari (see [3,15,16]). This BVP turns out to be well-posed locally (i.e. for small a) under assumptions which agree with the experimental set-up: in particular, estimates by defect of the crystal thickness a for which Cesari's existence and uniqueness theorem is valid can be computed and shown to be consistent with the thickness of crystals used in experiments to generate a second-harmonic wave [3-5]. Cesari's existence proof is constructive and yields a convergent iterative process which can be used to approximate the solution.

(iv) As soon as this BVP is solved, and the solution inside D_a is known, the reflected and transmitted waves can be completely determined via the continuous matching (6.86) with the solution inside the slab, exactly as in the linear isotropic case considered in §4.9. But, since here the Maxwell equations in D_a are nonlinear, the reflected and transmitted waves will include a small added term with frequency 2ω similar to (6.78), i.e. a second-harmonic wave (with period π/ω). The transmitted SH wave is particularly relevant, being the one that is usually detected in experiments. In fact, the second iteration step of Cesari's converging process for a quartz slab yields an approximate second-harmonic transmitted wave that is in full agreement with Franken and Ward's experimental results.

For further details the geometry and the crystal class must be specified. We consider below crystals of class 32-D3 and 6-C6.

6.7.3 Crystals of class 32-D3.

We first consider a uniaxial crystal of class 32-D3, like quartz [6]. We choose as coordinate axes (x_1, x_2, x_3) the principal axes (x, y, z) of the crystal, with z the optic axis (see sections 6.4 and 6.5). If

$$\boldsymbol{E} = \sum_{j=1}^{3} E_j \mathbf{c}_j \quad , \quad \boldsymbol{D} = \sum_{j=1}^{3} D_j \mathbf{c}_j \ , \quad \boldsymbol{H} = \sum_{j=1}^{3} H_j \mathbf{c}_j$$

the approximate nonlinear constitutive relations (6.80) for the vector $\boldsymbol{D}=\hat{\boldsymbol{\epsilon}}\boldsymbol{E}+\hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E}$ in the crystal have the form [6]

$$\begin{aligned} D_1 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_1 = \epsilon_\perp E_1 + \beta E_2 E_3 + 2\alpha E_1 E_2 \\ D_2 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_2 = \epsilon_\perp E_2 - \beta E_1 E_3 + \alpha(E_1^2 - E_2^2) \\ D_3 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_3 = \epsilon_\parallel E_3 \end{aligned} \tag{6.91}$$

[6] note that the axes x and y are interchanged here with respect to the usual convention adopted in [41] and in [23,24]

where the real constants α and β, which denote the sole non-zero entries of the tensor $\hat{\boldsymbol{\psi}}$, are very small in absolute value [41]. In the absence of dispersion, thermodynamic restrictions would imply that $\beta = 0$ [6, 17,20]. In the case of Franken & Ward's experiments for quartz the values which can be inferred from [41] for the relevant wavelength range are

$$\begin{aligned} &n_{or} \cong 1.54 \ , \ n_{ex} \cong 1.55 \ , \ |\alpha|/\epsilon_\perp \cong \ 0.2\ 10^{-12}\text{m/volt} \\ &E_l \cong 10^7\text{volt/m} \ , \ |\beta| \cong 0.025|\alpha| \end{aligned}$$

so that actually $|\beta| \ll |\alpha|$.

We examine separately the cases of propagation along the three principal axes.

6.7.4 Propagation along the $x-$axis.

Here $\boldsymbol{k}$=$\mathbf{c}_1$, $\boldsymbol{k}\cdot\boldsymbol{x}$= x, $D_a = \{0 \le x \le a\}$ and the incident wave is given by eqs. (6.87) with

$$\text{(6.92)} \qquad \boldsymbol{E}_o = E_l(\mathbf{c}_2 sin\,\theta + \mathbf{c}_3\,cos\,\theta) \quad , \boldsymbol{H}_o = \sqrt{\frac{\epsilon_o}{\mu_o}} E_l(\mathbf{c}_3 sin\,\theta - \mathbf{c}_2\,cos\,\theta)$$

where θ denotes the polarization angle, i.e. the angle between the electric field of the incident wave $\boldsymbol{E}_o$ and the optic axis z in the $(y, z)-$plane, and $E_l > 0$ is the electric field intensity of the laser beam. Eqs. (6.88), (89) and (6.91) imply that $H_1 \equiv D_1 \equiv 0$ so that $E_1 \equiv 0$ for $x < 0$, $x > a$ and

$$\epsilon_\perp E_1 + \beta E_2 E_3 + 2\alpha E_1 E_2 \equiv 0$$

in D_a, for all t. Therefore, for a generic value of θ, $\boldsymbol{E}$ and $\boldsymbol{H}$ are unknown vector functions of (x, t) of the form

$$\mathbf{E}(x,t) \ = \sum_{j=1}^{3} E_j(x,t)\,\mathbf{c}_j, \qquad \boldsymbol{H}(x,t) \ = H_2(x,t)\,\mathbf{c}_2 + H_3(x,t)\,\mathbf{c}_3$$

with E_1 vanishing outside D_a and satisfying the relation

$$\text{(6.93)} \qquad E_1 = -\frac{\beta E_2 E_3}{\epsilon_\perp + 2\alpha E_2}$$

in D_a, for all t.

The linear Maxwell equations (6.84), (6.85) outside the slab, i.e. for $x < 0$ and $x > a$, read

$$\epsilon_o \frac{\partial E_2}{\partial t} = -\frac{\partial H_3}{\partial x} \ , \ \mu_o \frac{\partial H_3}{\partial t} = -\frac{\partial E_2}{\partial x} \tag{6.94}$$

$$\epsilon_o \frac{\partial E_3}{\partial t} = \frac{\partial H_2}{\partial x} \ , \ \mu_o \frac{\partial H_2}{\partial t} = \frac{\partial E_3}{\partial x} \tag{6.95}$$

and thus split up into two separate blocks for (E_2, H_3) and (E_3, H_2) , respectively.

In the crystal slab ($0 \leq x \leq a$, $t \in \mathbb{R}$) the nonlinear Maxwell equations (6.82), (6.83) take the form

$$\begin{aligned} &\epsilon_\perp \frac{\partial E_2}{\partial t} + 2\alpha(E_1 \frac{\partial E_1}{\partial t} - E_2 \frac{\partial E_2}{\partial t}) - \beta(E_3 \frac{\partial E_1}{\partial t} + E_1 \frac{\partial E_3}{\partial t}) = -\frac{\partial H_3}{\partial x} \\ &\mu_o \frac{\partial H_3}{\partial t} = -\frac{\partial E_2}{\partial x} \end{aligned} \tag{6.96}$$

and

$$\epsilon_\| \frac{\partial E_3}{\partial t} = \frac{\partial H_2}{\partial x} \ , \ \mu_o \frac{\partial H_2}{\partial t} = \frac{\partial E_3}{\partial x} \tag{6.97}$$

where E_1 is given by eq. (6.93), and $H_1 \equiv 0$. Note that (6.97) is a linear system in the unknowns (E_3, H_2) independent of E_1, E_2 and H_3.

The general solution of (6.94) and of (6.95) outside the slab is given by the superposition of a progressive and a regressive traveling wave (see Chapter 4). The progressive wave in the first region $x < 0$ must coincide with the incident wave ($\boldsymbol{E}_{in}$,$\boldsymbol{H}_{\boldsymbol{in}}$), given by eqs. (6.87), (6.90) and (6.92) with $\boldsymbol{k}\cdot\boldsymbol{x} = x$

$$\begin{aligned} \boldsymbol{E}_{in} &= E_l(\mathbf{c}_2 sin\,\theta + \mathbf{c}_3\, cos\,\theta)\, cos(\omega(t - x/c_o)) \\ \boldsymbol{H}_{in} &= E_l\sqrt{\epsilon_o/\mu_o}\,(\mathbf{c}_3 sin\,\theta - \mathbf{c}_2\, cos\,\theta)\, cos(\omega(t - x/c_o)) \end{aligned}$$

and, as already mentioned, we assume that in the third region $x > a$ there is no regressive wave. If we denote by

$(\boldsymbol{E}^{\mathrm{I}}, \boldsymbol{H}^{\mathrm{I}})$ the e.m. field in the first region $x < 0$

$(\boldsymbol{E}^{\mathrm{III}}, \boldsymbol{H}^{\mathrm{III}})$ the e.m. field in the third region $x > a$

the form of the solution in the first region is given for all t by

$$
\begin{aligned}
E_3^{\mathrm{I}}(x,t) &= E_l\, cos(\omega(t - x/c_o))\, cos\,\theta + \mathcal{R}_3(t + x/c_o) \\
H_2^{\mathrm{I}}(x,t) &= \sqrt{\epsilon_o/\mu_o}\big[-E_l cos(\omega(t - x/c_o))\, cos\,\theta\; + \mathcal{R}_3(t + x/c_o)\big]
\end{aligned}
\tag{6.98}
$$

and

$$
\begin{aligned}
E_2^{\mathrm{I}}(x,t) &= E_l cos(\omega(t - x/c_o)) sin\,\theta\; + \mathcal{R}_2(t + x/c_o) \\
H_3^{\mathrm{I}}(x,t) &= \sqrt{\epsilon_o/\mu_o}\big[E_l\, cos(\omega(t - x/c_o)) sin\,\theta\; - \mathcal{R}_2(t + x/c_o)\big]
\end{aligned}
\tag{6.99}
$$

where $\mathcal{R}_j$ $(j = 2, 3)$ denote the components of the unknown reflected wave. In the third region $x > a$ the solution has the form

$$
\begin{aligned}
H_2^{\mathrm{III}}(x,t) &= -\sqrt{\epsilon_o/\mu_o} E_3^{\mathrm{III}}(x,t) = -\sqrt{\epsilon_o/\mu_o}\, \mathcal{T}_3(t - x/c_o) \\
H_3^{\mathrm{III}}(x,t) &= \sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{III}}(x,t) = \sqrt{\epsilon_o/\mu_o}\, \mathcal{T}_2(t - x/c_o)
\end{aligned}
\tag{6.100}
$$

where $\mathcal{T}_j$ $(j = 2, 3)$ denote the components of the unknown transmitted wave (Exercise 9). Moreover we have seen that

$$
E_1^{\mathrm{I}} = H_1^{\mathrm{I}} = E_1^{\mathrm{III}} = H_1^{\mathrm{III}} = 0 \tag{6.101}
$$

Thus, although $\mathcal{R}$ and $\mathcal{T}$ are unknown, the form of the solutions outside the crystal slab is known, and from this we immediately derive the equations

$$
\begin{aligned}
\sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{I}}(x,t) + H_3^{\mathrm{I}}(x,t) &= 2\sqrt{\epsilon_o/\mu_o} E_l\, cos(\omega(t - x/c_o))\, sin\,\theta \\
\sqrt{\epsilon_o/\mu_o} E_3^{\mathrm{I}}(x,t) - H_2^{\mathrm{I}}(x,t) &= 2\sqrt{\epsilon_o/\mu_o} E_l\, cos(\omega(t - x/c_o))\, cos\,\theta
\end{aligned}
$$

for all $x \leq 0, t \in \mathbb{R}$ and

$$
\begin{aligned}
\sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{III}}(x,t) - H_3^{\mathrm{III}}(x,t) &= 0 \\
\sqrt{\epsilon_o/\mu_o} E_3^{\mathrm{III}}(x,t) + H_2^{\mathrm{III}}(x,t) &= 0
\end{aligned}
$$

for all $x \geq a$, $t \in \mathbb{R}$. In particular taking $x = 0$ we see that the solution in the first region satisfies the linear inhomogeneous boundary condition

$$
\begin{aligned}
\sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{I}}(0,t) + H_3^{\mathrm{I}}(0,t) &= 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, sin\,\theta \\
\sqrt{\epsilon_o/\mu_o} E_3^{\mathrm{I}}(0,t) - H_2^{\mathrm{I}}(0,t) &= 2\sqrt{\epsilon_o/\mu_o} E_l\, cos(\omega t)\, cos\,\theta
\end{aligned}
$$

of impedance type for all $t \in \mathbb{R}$. Similarly, we see that the solution in the third region satisfies the linear homogeneous boundary condition for $x = a$

$$\sqrt{\epsilon_o/\mu_o}E_2^{\mathrm{III}}(a,t) - H_3^{\mathrm{III}}(a,t) = 0$$
$$\sqrt{\epsilon_o/\mu_o}E_3^{\mathrm{III}}(a,t) + H_2^{\mathrm{III}}(a,t) = 0$$

for all $t \in \mathbb{R}$. The continuous matching conditions (6.86) then imply that these impedance boundary conditions are satisfied also by the tangential components E_2, E_3, H_2, H_3 of the field inside the slab $0 \le x \le a$.

Summarizing, the components (E_2, H_3) of the field inside the crystal must satisfy the nonlinear Maxwell equations (6.96) for all $t \in \mathbb{R}$:

$$(6.102)\qquad \begin{aligned} -\frac{\partial H_3}{\partial x} &= \epsilon_\perp \frac{\partial E_2}{\partial t} + 2\alpha\,(E_1\frac{\partial E_1}{\partial t} - E_2\frac{\partial E_2}{\partial t}) - \beta\,(E_3\frac{\partial E_1}{\partial t} + E_1\frac{\partial E_3}{\partial t}) \\ -\frac{\partial E_2}{\partial x} &= \mu_o\frac{\partial H_3}{\partial t} \qquad (0 < x < a,\ t \in \mathbb{R}) \end{aligned}$$

together with the impedance boundary conditions on the slab walls

$$(6.103)\qquad \begin{aligned} \sqrt{\frac{\epsilon_o}{\mu_o}}E_2(0,t) + H_3(0,t) &= 2\sqrt{\frac{\epsilon_o}{\mu_o}}E_l\,cos\,\omega t\,sin\,\theta \\ \sqrt{\frac{\epsilon_o}{\mu_o}}E_2(a,t) - H_3(a,t) &= 0 \qquad (t \in \mathbb{R}) \end{aligned}$$

Similarly, the components (E_3, H_2) satisfy the linear Maxwell equations (6.97) in D_a for $t \in \mathbb{R}$:

$$(6.104)\qquad \epsilon_\parallel\frac{\partial E_3}{\partial t} = \frac{\partial H_2}{\partial x}\ ,\quad \mu_o\frac{\partial H_2}{\partial t} = \frac{\partial E_3}{\partial x} \qquad (0 < x < a,\ t \in \mathbb{R})$$

together with the impedance boundary conditions

$$(6.105)\qquad \begin{aligned} \sqrt{\frac{\epsilon_o}{\mu_o}}E_3(0,t) - H_2(0,t) &= 2\sqrt{\frac{\epsilon_o}{\mu_o}}E_l\,cos(\omega t)\,cos\,\theta \\ \sqrt{\frac{\epsilon_o}{\mu_o}}E_3(a,t) + H_2(a,t) &= 0, \qquad (t \in \mathbb{R}) \end{aligned}$$

As soon as the solution (E_2, E_3, H_2, H_3) in D_a is known, E_1 follows from eq. (6.93) and the reflected and transmitted waves follow from eqs. (6.99) and (6.100) :

$$(6.106)\qquad \mathcal{R}_2(t) = \frac{1}{2}\big[E_2(0,t) - \sqrt{\frac{\mu_o}{\epsilon_o}}H_3(0,t)\big]$$

$$\mathcal{R}_3(t) = \frac{1}{2}\big[E_3(0,t) + \sqrt{\frac{\mu_o}{\epsilon_o}} H_2(0,t)\big]$$

(6.107) $$\mathcal{T}_2(t) = E_2(a, t + a/c_o)\ ,\ \ \mathcal{T}_3(t) = E_3(a, t + a/c_o)$$

so that the solutions in the first and third regions ($\boldsymbol{E}^{\rm I}$,$\boldsymbol{H}^{\rm I}$), ($\boldsymbol{E}^{\rm III}$,$\boldsymbol{H}^{\rm III}$) are completely determined too.

Thus all we need is to solve the two boundary value problems (6.102)-(6.103) and (6.104)-(6.105).

We first consider the linear boundary value problem (6.104)-(6.105) for the field components E_3, H_2 in the slab. By virtue of Prroposition 6.7.1 below, this BVP has a unique bounded solution, and being linear its solution can be written in closed form. Let us define the refractive index $n_{ex} > 1$, reflection coefficient $0 < r_\| < 1$ and wavenumbers p_o, $p_\|$:

$$n_{ex} := \sqrt{\frac{\epsilon_\|}{\epsilon_o}}\ ,\ \ r_\| := \frac{n_{ex}-1}{n_{ex}+1}\ ,\ \ p_\| := \frac{\omega}{c_{ex}}\ ,\ \ c_{ex} := \frac{1}{\sqrt{\epsilon_\| \mu_o}}\ ,\ \ p_o = \frac{\omega}{c_o}$$

Moreover, let

(6.108) $$\begin{aligned} A_\| &:= 2E_l(n_{ex}+1)^{-1}\, d_\|^{-1}\big[1 - r_\|^2 exp(2ip_\| a)\big] \\ B_\| &:= r_\| A_\| exp(-2ip_\| a)\ ,\quad d_\| := 1 + r_\|^4 - 2r_\|^2 cos(2p_\| a) \end{aligned}$$

After some calculations we then find that (E_3, H_2) are given in D_a for all t by the real part of the complex expressions

(6.109) $$E_3(x,t) = Re\big[A_\| exp(i\omega t\ - ip_\| x) + B_\| exp(i\omega t + ip_\| x)\big] cos\,\theta$$

$$H_2(x,t) = \sqrt{\epsilon_\|/\mu_o}\, Re\big[-A_\| exp(i\omega t - ip_\| x) + B_\| exp(i\omega t + ip_\| x)\big] cos\,\theta$$

(Exercise 10). Eqs. (6.106) and (6.107) yield then

(6.110) $$\begin{aligned} \mathcal{R}_3(t) &= \frac{1}{2}(n_{ex}-1) Re\big[A_\|\, exp(i\omega t)(exp(-2ip_\| a) - 1)\big] cos\,\theta \\ \mathcal{T}_3(t) &= \ (1 + r_\|)\, Re\big[\,A_\| exp(i\omega t + ip_o a -\ ip_\| a)\big] cos\,\theta \end{aligned}$$

This linear solution is periodic with the same frequency ω as the incident laser wave, so that it cannot give rise to any second-harmonic term. Note that the $z-$component of the reflected wave $\mathcal{R}_3$ vanishes if $n_{ex} = 1$ or if $p_{\|}a$ is a multiple of π (cfr. §4.9). Clearly, the solutions $E_3(x,t)$ and $H_2(x,t)$ given by (6.108)-(6.110), are bounded, together with all their partial derivatives, for all $0 \le x \le a, t \in \mathbb{R}$, and periodic in t with the same period $2\pi/\omega$ of the incident wave. In particular we have

$$|E_3(x,t)| \le 2\frac{(1+|r_{\|}|)^3}{(1-r_{\|}^2)^2}E_l \ , \ |\frac{\partial E_3}{\partial t}| \le 2\,\omega\,\frac{(1+|r_{\|}|)^3}{(1-r_{\|}^2)^2}E_l \tag{6.111}$$

We consider now the nonlinear BVP (6.102)-(6.103) for (E_2, H_3) in the slab, where E_1 is defined by eq. (6.93) and E_3 by eq. (6.109). Let us put

$$E_2 = E_o + \tilde{E} \ , \quad H_3 = H_o + \tilde{H}$$

where E_o, H_o are the solutions of the linearized system, i.e. the system (6.102) with $\alpha = \beta = 0$

$$\begin{aligned} -\frac{\partial H_o}{\partial x} &= \epsilon_\perp \frac{\partial E_o}{\partial t} \\ -\frac{\partial E_o}{\partial x} &= \mu_o \frac{\partial H_o}{\partial t} \qquad (0 < x < a,\ t \in \mathbb{R}) \end{aligned} \tag{6.112}$$

together with the inhomogeneous impedance boundary conditions (6.103)

$$\begin{aligned} \sqrt{\frac{\epsilon_o}{\mu_o}}E_o(0,t) + H_o(0,t) &= \ 2\sqrt{\frac{\epsilon_o}{\mu_o}}\ E_l\, cos\,\omega t\, sin\,\theta \\ \sqrt{\frac{\epsilon_o}{\mu_o}}E_o(a,t) - H_o(a,t) &= \ 0 \quad (t \in \mathbb{R}) \end{aligned} \tag{6.113}$$

Again, the solution of this linear problem is periodic, with the same frequency ω as the incident laser wave, and can be written in closed form. Let us define the refractive index $n_{or} > 1$, reflection coefficient $0 < r_\perp < 1$ and wavenumber $p_\perp$

$$n_{or} := \sqrt{\frac{\epsilon_\perp}{\epsilon_o}}, \quad r_\perp := \frac{n_{or} - 1}{n_{or} + 1} \ , \ p_\perp := \frac{\omega}{c_{or}} \ , \ c_{or} := \frac{1}{\sqrt{\epsilon_\perp \mu_o}}$$

Moreover, let

$$\begin{aligned} A_\perp &:= 2E_l(n_{or}+1)^{-1}d_\perp^{-1}\big[1 - r_\perp^2 exp(2ip_\perp a)\big] \\ B_\perp &:= \ r_\perp A_\perp\, exp(-2ip_\perp a) \ , \ d_\perp := 1 + r_\perp^4 \ - 2r_\perp^2 cos(2p_\perp a) \end{aligned}$$

Proceeding like before we find that $(E_o,\, H_o)$ are given in D_a for all t by

$$\begin{aligned} E_o(x,t) &= Re\big[A_\perp exp(i\omega t\, - ip_\perp x) + B_\perp exp(i\omega t + ip_\perp x)\big]\, sin\,\theta \\ H_o(x,t) &= \sqrt{\epsilon_\perp/\mu_o}\, Re\big[A_\perp exp(i\omega t - ip_\perp x) \\ &\quad - B_\perp exp(i\omega t + ip_\perp x)\big]\, sin\theta \end{aligned} \tag{6.114}$$

and the linearized reflected and transmitted waves follow from eqs. (6.106) and (6.107)

$$\begin{aligned} \mathcal{R}_o(t) &:= \frac{1}{2}\big[E_o(0,t) - \sqrt{\frac{\mu_o}{\epsilon_o}}H_o(0,t)\big] \\ &= \frac{n_{or}-1}{2}\, Re\big[A_\perp(e^{-2ip_\perp a} - 1)e^{i\omega t}\big]\, sin\theta \end{aligned} \tag{6.115}$$

$$\mathcal{T}_o(t) := E_o(a, t{+}a/c_o) = (1{+}r_\perp)\, Re\big[\, A_\perp exp(i\omega t{+}ip_o a{-}\, ip_\perp a)\big]\, sin\theta \tag{6.116}$$

Note that $\mathcal{R}_o$ vanishes if either $n_{or} = 1$ or $p_\perp a$ is a multiple of π. The functions $E_o(x,t)$, $H_o(x,t)$ are bounded, together with all their partial derivatives, for all $0 \le x \le a, t \in \mathbb{R}$. In particular, we have

$$|E_o(x,t)| \le 2\frac{(1+|r_\perp|)^3}{(1-r_\perp^2)^2}E_l\ ,\ \ |H_o(x,t)| \le\, 2\frac{(1+|r_\perp|)^3}{(1-r_\perp^2)^2}\frac{\sqrt{\epsilon_\perp}}{\sqrt{\mu_o}}E_l \tag{6.117}$$

and

$$|\frac{\partial E_o}{\partial t}\,| \le 2\omega\, \frac{(1+|r_\perp|)^3}{(1-r_\perp^2)^2}E_l \tag{6.118}$$

for all $0 \le x \le a, t \in \mathbb{R}$.

We come now to the equations satisfied by the deviations $\tilde{E}$ and $\tilde{H}$. Let

$$\Lambda_1 := \beta\, E_3\ ,\ \ \Lambda_2 := \epsilon_\perp + 2\alpha(E_o + \tilde{E})$$

$$\eta(\tilde{E}) := 2\alpha\frac{\Lambda_1^2}{\Lambda_2^3}(E_o + \tilde{E}) - \frac{2\alpha}{\epsilon_\perp}\,(E_o + \tilde{E}) + \beta\frac{\Lambda_1}{\Lambda_2^2}E_3$$

and

$$\begin{aligned} \mathcal{F}(\tilde{E}) &:= 2\alpha\epsilon_\perp\frac{\Lambda_1^2}{\Lambda_2^3}\frac{\partial E_o}{\partial t}(E_o + \tilde{E}) + 2\alpha\beta\frac{\Lambda_1}{\Lambda_2^2}\frac{\partial E_3}{\partial t}(E_o + \tilde{E})^2 + \beta\epsilon_\perp\frac{\Lambda_1}{\Lambda_2^2}E_3\frac{\partial E_o}{\partial t} \\ &+ \beta^2\frac{1}{\Lambda_2}E_3\frac{\partial E_3}{\partial t}(E_o + \tilde{E}) + \beta\frac{\Lambda_1}{\Lambda_2}\frac{\partial E_3}{\partial t}(E_o + \tilde{E}) \end{aligned}$$

where E_3 and E_o are given by (6.109) and (6.114), respectively. Then by eqs. (6.93), (6.112), (6.113) and manipulations, the deviations $\tilde{E}$, $\tilde{H}$ are found to satisfy the nonlinear Maxwell equations in the slab

$$\begin{aligned} &\frac{\partial \tilde{H}}{\partial x} + \epsilon_\perp \big[1 + \ \eta(\tilde{E})\big] \frac{\partial \tilde{E}}{\partial t} \\ &= 2\alpha (E_o + \tilde{E}) \frac{\partial E_o}{\partial t} + \mathcal{F}(\tilde{E}), \\ &\frac{\partial \tilde{E}}{\partial x} + \mu_o \frac{\partial \tilde{H}}{\partial t} = 0 \quad (0 < x < a,\, t \in \mathbb{R}) \end{aligned} \tag{6.119}$$

together with the homogeneous impedance boundary conditions

$$\sqrt{\frac{\epsilon_o}{\mu_o}} \tilde{E}(0,t) + \tilde{H}(0,t) = \ 0 \ \ , \ \ \sqrt{\frac{\epsilon_o}{\mu_o}} \tilde{E}(a,t) - \tilde{H}(a,t) = \ 0 \qquad (t \in \mathbb{R}) \tag{6.120}$$

Since $|\beta| < |\alpha|$, we can define a single non-dimensional parameter

$$\delta := \frac{|\alpha|}{\epsilon_\perp} E_l$$

which is very small in all cases of interest ($\delta \cong 0.2\, 10^{-5}$ in [23,24]). We have then

$$\eta(\tilde{E}) = -\frac{2\alpha}{\epsilon_\perp} (E_o + \tilde{E}) + O(\delta^2) \ \ , \ \mathcal{F}(\tilde{E}) = O(\delta^2) \qquad \text{as } \delta \to 0 \tag{6.121}$$

Proposition 6.7.1 *If δ, a are sufficiently small, there exists a unique solution $(\tilde{E}, \tilde{H})$ of (6.119) and (6.120), bounded and Lipschitz continuous together with its first derivatives in $D_a \times \mathbb{R}$. This solution is periodic in t with period $2\pi/\omega$, and depends continuously on the incident wave (in the uniform norm).*

The proof of this proposition is given in the Appendix. The sufficient conditions on δ and a required by the proof are found to be consistent with the actual values of crystal thickness and laser beam intensities used in experiments, where typically $\delta \cong 10^{-5}$ and $a \cong 1mm$ or less. In the case of linear systems, like (6.104)-(6.105), the restrictions on δ and a obviously drop out.

Remark 7. It is well-known that quasilinear hyperbolic systems may develop shocks after a finite time [2,18]. For the system (6.119) the time and space variables are interchanged and $x > 0$ is the hyperbolic variable, so that shocks may occur after a finite $x-$interval, say at $x = a_s$, $0 < a_s < +\infty$ (and for some $t = t_s$). Thus one cannot expect the above proposition to hold in the large, for $a \geq a_s$. On the other hand, we recall that the present model makes sense anyway only for sufficiently small values of a. The same remark applies to Proposition 6.7.2 and Proposition 6.7.3 below.

We proceed to show that $\tilde{E}(x,t), \tilde{H}(x,t)$ contain a second-harmonic term with frequency 2ω, i.e. with period π/ω. In order to do this, we apply the usual perturbative method to (6.115) and (6.116): we expand $\tilde{E}(x,t)$ and $\tilde{H}(x,t)$ into formal power expansions in δ

$$\tilde{E}(x,t) = \delta E'(x,t) + \delta^2 E''(x,t) + ... , \quad \tilde{H}(x,t) = \delta H'(x,t) + \delta^2 H''(x,t) + ...$$

substitute into (6.119), (6.120) and keep only terms of order $O(\delta)$, using eqs. (6.111), (6.117), (6.118), (6.121) and the fact that $|\beta| < |\alpha|$. As a result we obtain the linear inhomogeneous system

$$\begin{aligned} \frac{\partial H'}{\partial x} + \epsilon_\perp \frac{\partial E'}{\partial t} &= 2\epsilon_\perp \frac{E_o}{E_l}\frac{\partial E_o}{\partial t} \\ \frac{\partial E'}{\partial x} + \mu_o \frac{\partial H'}{\partial t} &= 0 \quad (0 < x < a,\, t \in \mathbb{R}) \end{aligned} \tag{6.122}$$

with the homogeneous boundary conditions (6.120). The solution E', H' to this linear problem in the slab (which exists and is unique by force of Proposition 6.7.1) is given by

$$\begin{aligned} E' = E_l Re\Big[& e^{2i\omega t}\{k_o + k_1 e^{-2ip_\perp x} + k_2 e^{2ip_\perp x} \\ & + k_3 x e^{-2ip_\perp x} + k_4 x e^{2ip_\perp x}\}\Big] sin^2\theta \end{aligned} \tag{6.123}$$

$$\begin{aligned} H' = E_l \sqrt{\frac{\epsilon_\perp}{\mu_o}}\, Re\Big[& e^{2i\omega t}\{k_5 e^{-2ip_\perp x} + k_6 e^{2ip_\perp x} \\ & + k_3 x e^{-2ip_\perp x} - k_4 x e^{2ip_\perp x}\}\Big] sin^2\theta \end{aligned} \tag{6.124}$$

and the second-harmonic reflected and transmitted waves, which follows from eqs. (6.106), (6.107), (6.115) and (6.116), are given by

$$\tilde{\mathcal{R}}(t) = \delta E_l\, Re\Big[k_7\, e^{2i\omega t}\Big] sin^2\theta \;\; , \;\; \tilde{\mathcal{T}}(t) = \delta E_l Re\Big[k_8 e^{2i\omega t}\Big] sin^2\theta$$

where $k_o, ..., k_8$ are suitable complex constants, with $k_7 \neq 0$ and $k_8 \neq 0$ (see Exercise 11).

This perturbative approximation $\delta E'$, $\delta H'$ coincides, up to terms of order $O(\delta^2)$, with the second iterate of a converging iteration process which arises as part of the existence proof and, since δ is very small, it contains all the relevant information (see the Appendix). In the particular case of a quartz crystal slab of thickness up to $0.045\,mm$ (100-wave crystal) with $\theta = \pi/2$, the numerical results reported in [3] show that stability to order 10^{-4} is reached at the fourth iteration step with full agreement with the perturbative solution.

Summarizing, we have found the exact solution (6.109) for the components (E_3, H_2) in the three regions, and an approximate solution up to the order $O(\delta)$ for the components (E_2, H_3) given by

$$E_2^I(x,t) = E_l cos(\omega(t - x/c_o)) sin\,\theta + \mathcal{R}_2(t + x/c_o)$$
$$H_3^{\mathrm{I}}(x,t) = \sqrt{\epsilon_o/\mu_o}\big[E_l\, cos(\omega(t - x/c_o)) sin\,\theta\, - \mathcal{R}_2(t + x/c_o)\big]$$

in the first region $(x < 0)$, by

$$E_2(x,t) = E_o(x,t) + \delta E'(x,t) + O(\delta^2)$$
$$H_3(x,t) = H_o(x,t) + \delta H'(x,t) + O(\delta^2)$$

in the slab $D_a (0 \leq x \leq a)$, and by

$$H_3^{\mathrm{III}}(x,t) = \sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{III}}(x,t) = \sqrt{\epsilon_o/\mu_o}\, \mathcal{T}_2(t - x/c_o)$$

in the third region $(x > a)$. The reflected and transmitted waves

$$\mathcal{R}_2(t) = \mathcal{R}_o(t) + \tilde{\mathcal{R}}(t), \quad \mathcal{T}_2(t) = \mathcal{T}_o(t) + \tilde{\mathcal{T}}(t) \quad (t \in \mathbb{R})$$

are given by eqs. (6.115), (6.116), (6.123) and are related to the fields in the first and third regions by eqs. (6.98)−(6.100).

From this analysis we draw the following conclusions:

(i) The second-harmonic reflected and transmitted waves depend upon the polarization angle θ between the optic axis and the electric field of the incident laser wave, and they disappear if $\theta = 0$ or $\theta = \pi$.

(ii) The second-harmonic reflected and transmitted waves have intensities proportional to $sin^4\theta$, and are linearly polarized, with the electric field parallel to the y-axis and magnetic field parallel to the optic axis.

(iii) If the crystal thickness a is a multiple of $\pi/p_{\|}$ the reflected wave is polarized with the electric field orthogonal to the optic axis and disappears altogether if $\theta = 0$ or $\theta = \pi$.

(iv) The electromagnetic field is transversal everywhere with the exception of a small longitudinal component of the electric field in the crystal slab D_a, given by [7]

$$E_1 = -\frac{\beta E_o E_3}{\epsilon_\perp} + O(\delta^2)$$

All these conclusions are confirmed by the experiments [23,24].

6.7.5 Propagation along the $y-$axis.

Here $\boldsymbol{k}$=$\mathbf{c}_2$, $\boldsymbol{k}\cdot\boldsymbol{x}$= y, $D_a = \{0 \le y \le a\}$ and the incident wave is given by eq. (6.87) with

$$\boldsymbol{E}_o = E_l(\mathbf{c}_1 sin\,\theta + \mathbf{c}_3\, cos\,\theta)\ \ ,\ \ \boldsymbol{H}_o = \sqrt{\frac{\epsilon_o}{\mu_o}} E_l(\mathbf{c}_1\, cos\,\theta - \mathbf{c}_3 sin\,\theta)$$

where θ is the polarization angle, defined as the angle between the electric field of the incident wave $\boldsymbol{E}_o$ and the optic axis z in the $(x, z)-$plane. Eqs. (6.88), (6.89) and (6.91) yield $H_2 \equiv 0$ everywhere, $D_2 = \epsilon_o E_2 = 0$ for $y < 0$, $y > a$,and

$$D_2 \equiv \epsilon_\perp E_2 - \beta E_1 E_3 + \alpha(E_1^2 - E_2^2)\ = 0$$

in D_a, for all t. Therefore, for a generic value of θ, $\boldsymbol{E}$ and $\boldsymbol{H}$ are unknown vector functions of (y, t) of the form

$$\boldsymbol{E}(y, t) = \sum_{j=1}^{3} E_j(y, t)\, \mathbf{c}_j \qquad , \qquad \boldsymbol{H}(y, t) = H_1(y, t)\, \mathbf{c}_1 + H_3(y, t)\, \mathbf{c}_3$$

where E_2 is identically zero outside the crystal and is given by

$$E_2 = \frac{\epsilon_\perp}{2\alpha}\left[1 - \sqrt{1 - 4\alpha(\beta E_1 E_3 - \alpha E_1^2)/\epsilon_\perp^2}\ \right] \tag{6.125}$$

[7] more precisely, for dimensional homogeneity the remainder should be written as $E_l O(\varepsilon^2)$. A similar remark applies elsewhere

in the slab D_a for all t. The nonlinear Maxwell equations (6.82), (6.83) in the slab take the form

$$\epsilon_\perp \frac{\partial E_1}{\partial t} + 2\alpha(E_1 \frac{\partial E_2}{\partial t} + E_2 \frac{\partial E_1}{\partial t}) + \beta(E_2 \frac{\partial E_3}{\partial t} + E_3 \frac{\partial E_2}{\partial t}) = \frac{\partial H_3}{\partial y} \tag{6.126}$$
$$\mu_o \frac{\partial H_3}{\partial t} = \frac{\partial E_1}{\partial y}$$

and

$$\epsilon_\parallel \frac{\partial E_3}{\partial t} = -\frac{\partial H_1}{\partial y} \ , \quad \mu_o \frac{\partial H_1}{\partial t} = -\frac{\partial E_3}{\partial y} \tag{6.127}$$

where E_2 is given by eq. (6.125), and $H_2 \equiv 0$. Proceeding as in §6.7.4 we obtain the following set of impedance boundary conditions on the slab walls:

$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_3(0,t) + H_1(0,t) = 2\sqrt{\frac{\epsilon_o}{\mu_o}} E_l \cos(\omega t) \cos\theta \tag{6.128}$$
$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_3(a,t) - H_1(a,t) = 0 \qquad (t \in \mathbb{R})$$

and

$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_1(0,t) - H_3(0,t) = 2\sqrt{\frac{\epsilon_o}{\mu_o}}\ E_l \cos\omega t \sin\theta \tag{6.129}$$
$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_1(a,t) + H_3(a,t) = 0 \qquad (t \in \mathbb{R})$$

We first consider the linear boundary value problem (6.127), (6.128) for the field components E_3, H_1 in the slab, i.e. for $0 < y < a$, $t \in \mathbb{R}$. By virtue of Proposition 6.7.1 this BVP has a unique bounded solution (E_3, H_1), periodic in t with the same frequency ω as the incident laser wave, and given in closed form by an obvious adaptation of eq. (6.109). Hence the corresponding reflected and transmitted waves in the first and third regions are also periodic in t with frequency ω, so that they cannot contain any second-harmonic term.

We consider next the nonlinear BVP (6.126), (6.129) for (E_1, H_3) in the slab, where E_2 is defined by eq. (6.125) and E_3 is known from the previous step and satisfies boundedness estimates of the type (6.111). Let us put

$$E_1 = E_o + \tilde{E} \qquad , \quad H_3 = H_o + \tilde{H}$$

and let E_o, H_o satisfy the linearized system

$$\epsilon_\perp \frac{\partial E_o}{\partial t} = \frac{\partial H_o}{\partial y} \quad , \qquad \mu_o \frac{\partial H_o}{\partial t} = \frac{\partial E_o}{\partial y} \tag{6.130}$$

for $0 < x < a$, $t \in \mathbb{R}$, together with the inhomogeneous boundary conditions

$$\begin{aligned} \sqrt{\frac{\epsilon_o}{\mu_o}} E_o(0,t) - H_o(0,t) &= 2\sqrt{\frac{\epsilon_o}{\mu_o}}\; E_l \, cos\, \omega t \, sin\, \theta \\ \sqrt{\frac{\epsilon_o}{\mu_o}} E_o(a,t) + H_o(a,t) &= 0 \end{aligned} \tag{6.131}$$

for $t \in \mathbb{R}$. This linear solution is periodic, with the same frequency ω as the incident laser wave, and can be written in closed form by an obvious adaptation of eq. (6.114).

A second-harmonic term with frequency 2ω may arise only from the fields $(\tilde{E}, \tilde{H})$, which satisfy the system

$$\begin{aligned} &\frac{\partial \tilde{H}}{\partial y} - \epsilon_\perp \frac{\partial \tilde{E}}{\partial t} - 2\alpha(E_1 \frac{\partial E_2}{\partial t} + E_2 \frac{\partial E_1}{\partial t}) - \beta(E_2 \frac{\partial E_3}{\partial t} + E_3 \frac{\partial E_2}{\partial t}) = 0 \\ &\frac{\partial \tilde{E}}{\partial y} - \mu_o \frac{\partial \tilde{H}}{\partial t} = 0 \end{aligned}$$

in which $E_1 = E_o + \tilde{E}$, E_3 satisfies (6.111) and E_2 is given by eq. (6.125).By force of eq. (6.111) we can write

$$\frac{\partial E_2}{\partial t} = \eta_1(\tilde{E}) + \eta_2(\tilde{E}) \frac{\partial \tilde{E}}{\partial t}$$

Let

$$\delta := \frac{|\alpha|}{\epsilon_\perp} E_l \tag{6.132}$$

denote the small non-dimensional parameter defined previously. Then

$$E_2, \;\; \eta_1(\tilde{E}), \, \eta_2(\tilde{E}) = O(\delta)$$

as $\delta \to 0$.

Summarizing, the field $\tilde{E}$, $\tilde{H}$ must satisfy the nonlinear boundary value problem in the slab

$$\begin{aligned} &\frac{\partial \tilde{H}}{\partial y} - \epsilon_\perp \big[1+\eta(\tilde{E})\big] \frac{\partial \tilde{E}}{\partial t} = \mathcal{F}(\tilde{E}) \ , \\ &\frac{\partial \tilde{E}}{\partial y} \ = \mu_o \frac{\partial \tilde{H}}{\partial t} \ \ (0 < y < a, \, t \in \mathbb{R}) \end{aligned} \tag{6.133}$$

$$\sqrt{\frac{\epsilon_o}{\mu_o}}\tilde{E}(0,t) = \tilde{H}(0,t) \quad , \quad \sqrt{\frac{\epsilon_o}{\mu_o}}\tilde{E}(a,t) + \tilde{H}(a,t) = \ 0 \quad (t \in \mathbb{R}) \tag{6.134}$$

with $\eta(\tilde{E})$ and $\mathcal{F}(\tilde{E})$ of order δ^2

$$\eta(\tilde{E}) \, , \, \mathcal{F}(\tilde{E}) = O(\delta^2)$$

as $\delta \to 0$ [6]. The perturbative approximation to $\tilde{E}, \tilde{H}$ is obtained by putting

$$\tilde{E}(y,t) = \delta E'(y,t) + \delta^2 E''(\,y,t) + ... \ \ , \quad \tilde{H}(y,t) = \delta H'(y,t) + \delta^2 H''(y,t) + ...$$

substituting into (6.133), (6.134), and keeping only terms of order $O(\delta)$. In this way we obtain the linear homogeneous system for E', H'

$$\frac{\partial H'}{\partial y} - \epsilon_\perp \frac{\partial E'}{\partial t} = 0 \, , \ \ \frac{\partial E'}{\partial y} \ - \ \mu_o \frac{\partial H'}{\partial t} = 0 \qquad (0 < y < a, \, t \in \mathbb{R})$$

which coupled with the homogeneous boundary conditions (6.134) implies, by Proposition 6.7.1, that

$$E'(y,t) \equiv H'(y,t) \equiv 0$$

for all $0 \le y \le a, \, t \in \mathbb{R}$. Hence the corresponding approximations for the reflected and transmitted waves vanish, too. As already mentioned, the perturbative solution $\delta E'$, $\delta H'$ coincides up to terms of order $O(\delta^2)$ with the first iterate of the converging iteration process arising from the existence proof. Therefore the solution $(\tilde{E}, \tilde{H})$ to (6.133), (6.134) is of order δ^2 :

$$\tilde{E}(y,t) = \ O(\delta^2) \ , \ \ \tilde{H}(y,t) = O(\delta^2)$$

the solution in the slab is

$$E_2 = E_o + O(\delta^2) \quad , \qquad H_3(x,t) = H_o(x,t) + O(\delta^2)$$

and the corresponding reflected and transmitted waves in the first and third regions are of the form

$$\mathcal{R}_1(t) = \mathcal{R}_o(t) + O(\delta^2), \quad \mathcal{T}_1(t) = \mathcal{T}_o(t) + O(\delta^2) \quad (t \in \mathbb{R})$$

with $\mathcal{R}_o(t)$, $\mathcal{T}_o(t)$ periodic in t with frequency ω (Exercise 12).

We conclude that the second-harmonic wave is (at most) of order δ^2, and hence is too small to be observed in practice.

This conclusion is confirmed by the experiments, which show that no second-harmonic generation occurs for propagation along the $y-$axis [8] of a quartz crystal.

6.7.6 Propagation along the optic axis.

Here $\boldsymbol{k}=\mathbf{c}_3$, $\boldsymbol{k}\cdot\boldsymbol{x}= z$, $D_a = \left\{0 \le z \le a\right\}$ and the incident wave is given by eq. (6.87) with

$$(6.135) \quad \boldsymbol{E}_o = E_l(\mathbf{c}_1 sin\,\theta + \mathbf{c}_2\, cos\,\theta) \quad , \ \boldsymbol{H}_o = \sqrt{\frac{\epsilon_o}{\mu_o}} E_l(-\mathbf{c}_1\, cos\,\theta + \mathbf{c}_2 sin\,\theta)$$

where the polarization angle θ is now defined as the angle between the $y-$axis and the electric field of the incident laser wave. Eqs. (6.88), (6.89) and (6.91) yield $D_3 = \epsilon_o E_3$ for $z < 0$ and $z > a$, $D_3 \equiv \epsilon_{\|} E_3$ for $0 \le z \le a$ and hence

$$H_3 \equiv D_3 \equiv E_3 \equiv 0$$

everywhere. Therefore the waves are transversal everywhere, and, for a generic value of θ, $\boldsymbol{E}$ and $\boldsymbol{H}$ are unknown vector functions of (z,t) of the form

$$\boldsymbol{E}(z,t) \ = E_1(z,t)\,\mathbf{c}_1 + E_2(z,t)\,\mathbf{c}_2 \quad , \quad \boldsymbol{H}(z,t) = H_1(z,t)\,\mathbf{c}_1 + H_2(z,t)\,\mathbf{c}_2$$

The linear Maxwell equations (6.84), (6.85) outside the slab i.e. for $z < 0$ and $z > a$, read

$$\epsilon_o \frac{\partial E_1}{\partial t} = -\frac{\partial H_2}{\partial z} \ , \ \mu_o \frac{\partial H_2}{\partial t} = -\frac{\partial E_1}{\partial z}$$
$$\epsilon_o \frac{\partial E_2}{\partial t} = \frac{\partial H_1}{\partial z} \ , \ \mu_o \frac{\partial H_1}{\partial t} = \frac{\partial E_2}{\partial z}$$

[8] $x-$axis in the notations of [23,24]

The solution of this system can be obtained from eqs. (6.98) and ff. by straightforward substitutions. For $z < 0$ we have:

$$\begin{aligned} E_2^{\mathrm{I}}(x,t) &= E_l\, cos(\omega(t - z/c_o))\, cos\,\theta + \mathcal{R}_2(t + z/c_o) \\ H_1^{\mathrm{I}}(z,t) &= \sqrt{\epsilon_o/\mu_o}\big[-E_l\, cos(\omega(t - z/c_o))\, cos\,\theta \; + \mathcal{R}_2(t + z/c_o)\big] \end{aligned} \tag{6.136}$$

and

$$\begin{aligned} E_1^{\mathrm{I}}(z,t) &= E_l cos(\omega(t - z/c_o)) sin\,\theta \; + \mathcal{R}_1(t + z/c_o) \\ H_2^{\mathrm{I}}(x,t) &= \sqrt{\epsilon_o/\mu_o}\big[E_l\, cos(\omega(t - z/c_o)) sin\,\theta \; - \mathcal{R}_1(t + z/c_o)\big] \end{aligned} \tag{6.137}$$

where $\mathcal{R}_j$ $(j = 1, 2)$ denote the components of the unknown reflected wave. In the third region $z > a$ the solution has the form

$$\begin{aligned} H_1^{\mathrm{III}}(z,t) &= -\sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{III}}(z,t) = -\sqrt{\epsilon_o/\mu_o}\, \mathcal{T}_2(t - z/c_o) \\ H_2^{\mathrm{III}}(z,t) &= \sqrt{\epsilon_o/\mu_o} E_1^{\mathrm{III}}(z,t) = \sqrt{\epsilon_o/\mu_o}\, \mathcal{T}_1(t - z/c_o) \end{aligned} \tag{6.138}$$

where $\mathcal{T}_j$ $(j = 1, 2)$ denote the components of the unknown transmitted wave. By sum and subtraction we obtain from (6.136) and (6.137) for $z = 0$ the relations

$$\begin{aligned} \sqrt{\epsilon_o/\mu_o} E_1^{\mathrm{I}}(0,t) + H_2^{\mathrm{I}}(0,t) &= 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, sin\,\theta \\ \sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{I}}(0,t) - H_1^{\mathrm{I}}(0,t) &= 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, cos\,\theta \end{aligned} \tag{6.139}$$

and

$$\begin{aligned} \mathcal{R}_1(t) &= \frac{1}{2}\big[E_1^{\mathrm{I}}(0,t) - \sqrt{\mu_o/\epsilon_o}\, H_2^{\mathrm{I}}(0,t)\big] \\ \mathcal{R}_2(t) &= \frac{1}{2}\big[E_2^{\mathrm{I}}(0,t) + \sqrt{\mu_o/\epsilon_o}\, H_1^{\mathrm{I}}(0,t)\big] \end{aligned} \tag{6.140}$$

Similarly, eq. (6.138) for $z = a$ yields

$$\begin{aligned} \sqrt{\epsilon_o/\mu_o} E_1^{\mathrm{III}}(a,t) - H_2^{\mathrm{III}}(a,t) &= 0 \\ \sqrt{\epsilon_o/\mu_o} E_2^{\mathrm{III}}(a,t) + H_1^{\mathrm{III}}(a,t) &= 0 \end{aligned} \tag{6.141}$$

and

$$\mathcal{T}_1(t) = E_1^{\mathrm{III}}(a, t + a/c_o) \;\;,\;\; \mathcal{T}_2(t) = E_2^{\mathrm{III}}(a, t + a/c_o) \tag{6.142}$$

for all $t \in \mathbb{R}$. The continuous matching conditions (6.86) imply that the impedance boundary conditions (6.139), (6.141) are satisfied also by the tangential components E_1, E_2, H_1, H_2 of the field inside the slab:

$$\begin{aligned} &\sqrt{\epsilon_o/\mu_o}E_1(0,t) + H_2(0,t) = 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, sin\,\theta \\ &\sqrt{\epsilon_o/\mu_o}E_2(0,t) - H_1(0,t) = 2\sqrt{\epsilon_o/\mu_o}E_l\, cos(\omega t)\, cos\,\theta \\ &\sqrt{\epsilon_o/\mu_o}E_1(a,t) = H_2(a,t)\ ,\ \ \sqrt{\epsilon_o/\mu_o}E_2(a,t) + H_1(a,t) = 0 \end{aligned} \tag{6.143}$$

In addition, eqs. (6.82), (6.83) yield the nonlinear system in the four unknowns E_1, E_2, H_1, H_2 in D_a for all t

$$\begin{aligned} &(\epsilon_\perp + 2\alpha E_2)\frac{\partial E_1}{\partial t} + 2\alpha E_1 \frac{\partial E_2}{\partial t} = -\frac{\partial H_2}{\partial z}\ ,\ \ \mu_o\frac{\partial H_2}{\partial t} = -\frac{\partial E_1}{\partial z} \\ &2\alpha E_1\frac{\partial E_1}{\partial t} + (\epsilon_\perp - 2\alpha E_2)\frac{\partial E_2}{\partial t} = \frac{\partial H_1}{\partial z}\ \ ,\ \ \mu_o\frac{\partial H_1}{\partial t} = \frac{\partial E_2}{\partial z} \end{aligned} \tag{6.144}$$

The reflected and transmitted waves then follow from eqs. (6.140) and (6.142) by the continuous matching at $z = 0, a$

$$\mathcal{R}_1(t) = \frac{1}{2}\big[E_1(0,t) - \sqrt{\frac{\mu_o}{\epsilon_o}}\, H_2(0,t)\big]\ ,\ \ \mathcal{R}_2(t) = \frac{1}{2}\big[E_2(0,t) + \sqrt{\frac{\mu_o}{\epsilon_o}}\, H_1(0,t)\big]$$
$$\mathcal{T}_1(t) = E_1(z, t + a/c_o)\ \ ,\ \ \mathcal{T}_2(t) = E_2(z, t + a/c_o)$$

so that the solutions in the first and third regions $(\boldsymbol{E}^{\mathrm{I}},\boldsymbol{H}^{\mathrm{I}})$, $(\boldsymbol{E}^{\mathrm{III}},\boldsymbol{H}^{\mathrm{III}})$ are completely determined by eqs. (6.136)-(6.138) as soon as the solution inside the slab is known.

Unfortunately, Cesari's theory does not apply to the nonlinear problem (6.143), (6.144) except for particular values of the polarization angle θ. The reason is that for a generic value of θ there seems to be no way to reduce the nonlinear hyperbolic system (6.144) in four unknowns to characteristic form with a diagonally dominant matrix as required (see the Appendix). The situation changes if $\theta = k\pi/3$ $(k = 0, 1, 2, ...)$, since then one can prove the existence of an exact solution to (6.143), (6.144).

Proposition 6.7.2 *If δ, a are sufficiently small, and if*

$$tan\theta = -tan2\theta$$

there exists a solution (E_1, E_2, H_1, H_2) of (6.143) and (6.144) with the properties specified in Proposition 6.7.1.

Proof. If $\theta = 0, \pi$ then $sin\theta = 0$ and (6.143), (6.144) are satisfied exactly by taking $E_1 \equiv H_2 \equiv 0$ and E_2, H_1 as the unique solutions to the nonlinear 2×2 boundary value problem

$$\frac{\partial H_1}{\partial z} = (\epsilon_\perp - 2\alpha E_2)\frac{\partial E_2}{\partial t} \ , \ \frac{\partial E_2}{\partial z} = \mu_o \frac{\partial H_1}{\partial t}$$
$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_2(0,t) - H_1(0,t) = \pm 2\sqrt{\frac{\epsilon_o}{\mu_o}} E_l\, cos(\omega t) \ , \ \sqrt{\frac{\epsilon_o}{\mu_o}} E_2(a,t) + H_1(a,t) = 0$$

to which Cesari's theory does apply. Similarly, if $\theta = \pi/3, 2\pi/3, 4\pi/3, 5\pi/3$ then $tan\theta = \pm\sqrt{3}$ and an exact solution is given by

$$E_1 = \pm\sqrt{3}\, E_2 \ , \ H_2 = \mp\sqrt{3}\, H_1$$

with E_2, H_1 the unique solutions to the nonlinear 2×2 problem

$$\frac{\partial H_1}{\partial z} = (\epsilon_\perp + 4\alpha E_2)\frac{\partial E_2}{\partial t} \ , \ \frac{\partial E_2}{\partial z} = \mu_o \frac{\partial H_1}{\partial t}$$
$$\sqrt{\frac{\epsilon_o}{\mu_o}} E_2(0,t) - H_1(0,t) = 2\sqrt{\frac{\epsilon_o}{\mu_o}} E_l\, cos(\omega t)\, cos\theta_o \ , \ \sqrt{\frac{\epsilon_o}{\mu_o}} E_2(a,t) + H_1(a,t) = 0$$

to which again Cesari's theory does apply. This concludes the proof.

Note that one cannot exclude the existence of other, different exact solutions of (6.143) and (6.144), and therefore the solution whose existence is guaranteed by Proposition 6.7.2 might (in principle) not be unique.

A rigorous mathematical theory is therefore not available in the case of propagation parallel to the optic axis for a generic polarization of the incident wave. This implies that the following perturbation analysis will lack a rigorous mathematical interpretation. Let us define the deviations

$$\tilde{E}_1 := E_1 - E_1^o \ , \ \tilde{H}_1 := H_1 - H_1^o$$
$$\tilde{E}_2 := E_2 - E_2^o \ , \ \tilde{H}_2 := H_2 - H_2^o$$

If $(E_1^o, E_2^o, H_1^o, H_2^o)$ is the linearized solution in the crystal, satisfying the boundary conditions (6.143) and the linearized system

$$\epsilon_\perp \frac{\partial E_1}{\partial t} = -\frac{\partial H_2}{\partial z} \ , \ \mu_o \frac{\partial H_2}{\partial t} = -\frac{\partial E_1}{\partial z}$$
$$\epsilon_\perp \frac{\partial E_2}{\partial t} = \frac{\partial H_1}{\partial z} \ , \ \mu_o \frac{\partial H_1}{\partial t} = \frac{\partial E_2}{\partial z} \tag{6.145}$$

the deviations satisfy the nonlinear system

$$\frac{\partial \tilde{E}_1}{\partial z} + \mu_o \frac{\partial \tilde{H}_2}{\partial t} = 0 \ , \ \frac{\partial \tilde{H}_2}{\partial z} + (\epsilon_\perp + \eta_2)\frac{\partial \tilde{E}_1}{\partial t} + \eta_1 \frac{\partial \tilde{E}_2}{\partial t} = \mathcal{F}_1$$
$$\frac{\partial \tilde{E}_2}{\partial z} - \mu_o \frac{\partial \tilde{H}_1}{\partial t} = 0 \ , \ \frac{\partial \tilde{H}_1}{\partial z} - \eta_1 \frac{\partial \tilde{E}_1}{\partial t} - (\epsilon_\perp - \eta_2)\frac{\partial \tilde{E}_2}{\partial t} = \mathcal{F}_2 \tag{6.146}$$

and the homogeneous boundary conditions

$$\sqrt{\epsilon_o/\mu_o}\,\tilde{E}_1(0,t) + \tilde{H}_2(0,t) = 0 \ , \ \sqrt{\epsilon_o/\mu_o}\,\tilde{E}_2(0,t) - \tilde{H}_1(0,t) = 0$$
$$\sqrt{\epsilon_o/\mu_o}\tilde{E}_1((a,t) - \tilde{H}_2(a,t) = 0 \ , \ \sqrt{\epsilon_o/\mu_o}\,\tilde{E}_2(a,t) + \tilde{H}_1(a,t) = 0 \tag{6.147}$$

where

$$\eta_j := 2\alpha(E_j^o + \tilde{E}_j) \qquad (j = 1, 2) \tag{6.148}$$

and $\mathcal{F}_j = \mathcal{F}_j(\boldsymbol{E}^o, \tilde{\boldsymbol{E}})$ are linear functions of $\tilde{\boldsymbol{E}} = (\tilde{E}_1, \tilde{E}_2)$ whose explicit expressions can be easily written down. Cesari's theory does apply to the linearized problem (6.143)−(6.145), and it is easy to see that the linearized solution must have the form

$$E_1^o = E_o(z,t) sin\,\theta \ , \quad H_2^o = -H_o(z,t) sin\,\theta$$

and

$$E_2^o = E_o(z,t)\, cos\,\theta \quad , \quad H_1^o = H_o(z,t)\, cos\,\theta$$

where $E_o(z,t)$, $H_o(z,t)$ are combinations of sinusoidal functions, of period $2\pi/\omega$ (Exercise 13). The formal perturbation expansions for the nonlinear problem (6.146), (6.147) in terms of the small parameter δ defined by (6.122) reads

$$\tilde{E}_1 = \delta E_1'(z,t) + O(\delta^2) \ , \ \tilde{H}_2 = \delta H_2'(z,t) + O(\delta^2)$$
$$\tilde{E}_2 = \delta E_2'(z,t) + O(\delta^2) \ , \ \tilde{H}_1 = \delta H_1'(z,t) + O(\delta^2)$$

and by keeping only terms of order $O(\delta)$ we find the inhomogeneous linear system

$$\epsilon_\perp \frac{\partial E_1'}{\partial t} + \frac{\partial H_2'}{\partial z} = -\frac{2\alpha}{\delta}(E_1^o \frac{\partial E_2^o}{\partial t} + E_2^o \frac{\partial E_1^o}{\partial t}) \ , \ \mu_o \frac{\partial H_2'}{\partial t} + \frac{\partial E_1'}{\partial z} = 0$$
$$\epsilon_\perp \frac{\partial E_2'}{\partial t} - \frac{\partial H_1'}{\partial z} = \frac{2\alpha}{\delta}(E_2^o \frac{\partial E_2^o}{\partial t} - E_1^o \frac{\partial E_1^o}{\partial t}) \ , \ \mu_o \frac{\partial H_1'}{\partial t} - \frac{\partial E_2'}{\partial z} = 0 \tag{6.149}$$

and the homogeneous boundary conditions

$$\sqrt{\epsilon_o/\mu_o}E_1'(0,t) + H_2'(0,t) = 0\,,\ \sqrt{\epsilon_o/\mu_o}E_2'(0,t) - H_1'(0,t) = 0 \tag{6.150}$$
$$\sqrt{\epsilon_o/\mu_o}E_1'(a,t) - H_2'(a,t) = 0\ ,\ \sqrt{\epsilon_o/\mu_o}E_2'(a,t) + H_1'(a,t) = 0$$

The previous equations imply that

$$E_1^o\frac{\partial E_2^o}{\partial t} + E_2^o\frac{\partial E_1^o}{\partial t} = E^o\,\frac{\partial E^o}{\partial t} sin2\theta,\ E_2^o\frac{\partial E_2^o}{\partial t} - E_1^o\frac{\partial E_1^o}{\partial t} = E^o\,\frac{\partial E^o}{\partial t}\,cos2\theta$$

where $E^o\partial E^o/\partial t$ is periodic of period π/ω (see Exercise 13). The system (6.149) and the boundary conditions (6.150) are then satisfied by putting

$$E_1' = -E'(z,t)\,sin2\theta\ ,\ \ H_2' = H'(z,t) sin2\theta \tag{6.151}$$
$$E_2' = E'(z,t)\,cos2\theta \quad H_1' = H'(z,t)\,cos2\theta$$

where $E'(z,t)$, $H'(z,t)$ are the solutions of the boundary value problem

$$\epsilon_\perp\frac{\partial E'}{\partial t} - \frac{\partial H'}{\partial z} = \frac{2\epsilon_\perp}{E_l}E^o\,\frac{\partial E^o}{\partial t}\ ,\ \mu_o\frac{\partial H'}{\partial t} - \frac{\partial E'}{\partial z} = 0 \quad (0 \le z \le a,\ t \in \mathbb{R})$$
$$\sqrt{\frac{\epsilon_o}{\mu_o}}E'(0,t) - H'(0,t) = 0\ ,\ \sqrt{\frac{\epsilon_o}{\mu_o}}E'(a,t) + H'(a,t) = 0 \qquad (t \in \mathbb{R})$$

which exist and are unique by force of Proposition 6.7.1. The second-harmonic reflected and transmitted waves for $z < 0$ and $z > a$, respectively, are then given by:

$$\tilde{\mathcal{R}}_1(t) = -\frac{\delta}{2}\big[E'(0,t) + \sqrt{\frac{\mu_o}{\epsilon_o}}\,H'(0,t)\big]sin2\theta\,,\ \tilde{\mathcal{T}}_1(t) = -\delta E'(a,t+a/c_o) sin2\theta$$
$$\tilde{\mathcal{R}}_2(t) = \frac{\delta}{2}\big[E'(0,t) + \sqrt{\frac{\mu_o}{\epsilon_o}}\,H'(0,t)\big]cos2\theta,\quad \tilde{\mathcal{T}}_2(t) = \delta E'(a,t+a/c_o)\,cos2\theta$$

(see eqs. (6.148)). These equations taken in conjunction with eqs. (6.151) imply that the intensities of the perturbations

$$\sqrt{|E_1'|^2 + |E_2'|^2} = |E'(z,t)|\ ,\ \ \sqrt{|H_1'|^2 + |H_2'|^2} = |H'(z,t)|$$

and of the second harmonic reflected and transmitted waves

$$\sqrt{|\tilde{\mathcal{R}}_1(t)|^2 + |\tilde{\mathcal{R}}_2(t)|^2} = \frac{1}{2}|\delta|\big|E'(0,t) + \sqrt{\frac{\mu_o}{\epsilon_o}}\,H'(0,t)\big|$$
$$\sqrt{|\tilde{\mathcal{T}}_1(t)|^2 + |\tilde{\mathcal{T}}_2(t)|^2} = |\delta|\big|E'(a,t+a/c_o)\big|$$

are independent of θ. The details are left as an exercise (Exercise 14).

Summarizing, on the basis of this perturbation analysis we conclude that

(i) The second-harmonic reflected and transmitted waves have intensities independent of the polarization angle of the incident wave θ (the angle between the electric field and the optic axis z).

(ii) The second-harmonic reflected and transmitted waves are linearly polarized, with polarization angle 2θ.

(iii) The polarizations of the incident and second-harmonic waves coincide if θ is any multiple of 60° (see Proposition 6.7.2), since then $sin\theta = \pm sin 2\theta$.

(iv) The electromagnetic field is independent of the constant β.

All these conclusions are confirmed by the experiments.

It therefore appears that the second-harmonic wave can always be described by the perturbative solution, even though the latter is not justified mathematically in the case of propagation along the optic axis.

6.7.7 Crystals of class 6-D6.

We now briefly consider the case of (uniaxial) crystals of class 6-C6 [7]. A typical crystal of this class is $LiIO_3$ (Lithium Iodate) [41]. Choosing principal axes $(x_1, x_2, x_3) = (x, y, z)$ as before, with z the optic axis, the approximate nonlinear constitutive relations (6.80) for the vector $\boldsymbol{D} = \hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E}$ in the crystal are

$$\begin{aligned} D_1 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_1 = \epsilon_\perp E_1 + 2\beta_1 E_2 E_3 + 2\alpha_1 E_1 E_3 \\ D_2 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_2 = \epsilon_\perp E_2 + 2\alpha_1 E_2 E_3 - 2\beta_1 E_1 E_3 \\ D_3 &:= (\hat{\boldsymbol{\epsilon}}\boldsymbol{E} + \hat{\boldsymbol{\psi}}\boldsymbol{E}\boldsymbol{E})_3 = \epsilon_\parallel E_3 + \alpha_2(E_1^2 + E_2^2) + \beta_2 E_3^2 \end{aligned} \tag{6.152}$$

where the real constants α_j and $\beta_j (j = 1, 2)$ denote the sole non-zero entries of the tensor $\hat{\boldsymbol{\psi}}$ and are small (in absolute value). In the absence of dispersion, thermodynamic restrictions would imply that $\beta_1 = 0$ and $\alpha_1 = \alpha_2$. In the case of a Lithium Iodate crystal the values which can be inferred from [41] for the relevant wavelength range are

$$n_{or} \cong 1.9 \ , \ n_{ex} \cong 1.75 \ , \ \alpha_1 \cong \alpha_2 \cong \alpha$$

and

$$\frac{|\beta_1|}{\epsilon_\perp} \cong 4\,10^{-13} m/volt, \quad \frac{|\beta_2|}{\epsilon_\perp} \cong 5\,10^{-12} m/volt \quad , \frac{|\alpha\,|}{\epsilon_\perp} \cong 10^{-11} m/volt$$

If we define here the non-dimensional parameter

$$\delta := \frac{E_l}{\epsilon_\perp} sup(|\beta_1|, |\beta_2|, |\alpha_1|\,, |\alpha_2|) \tag{6.153}$$

then in all relevant cases δ is very small (of order 10^{-5} for $LiIO_3$ in typical experiments). Since usually, as we said above,

$$|\beta_1| \ll |\beta_2| < |\alpha_1| = |\alpha_2| := |\alpha|$$

he parameter δ can also be defined formally as in (6.132).

6.7.8 Propagation along the $x-$axis.

In this case, existence and uniqueness can be proven by adapting the arguments of the previous subsections. The first relation (6.152) yields

$$\epsilon_\perp E_1 + 2\beta_1 E_2 E_3 + 2\alpha_1 E_1 E_3 = 0$$

so that the longitudinal electric field E_1 in the slab is of order δ. Taking $E_1 = 0$ for simplicity, the nonlinear Maxwell equations in the slab are

$$\begin{aligned} -\frac{\partial H_3}{\partial x} &= \epsilon_\perp \frac{\partial E_2}{\partial t} + 2\alpha_1 (E_2 \frac{\partial E_3}{\partial t} + E_3 \frac{\partial E_2}{\partial t}) \\ -\frac{\partial E_2}{\partial x} &= \mu_o \frac{\partial H_3}{\partial t} \ , \ \ \frac{\partial E_3}{\partial x} = \mu_o \frac{\partial H_2}{\partial t} \\ \frac{\partial H_2}{\partial x} &= \ \epsilon_\| \frac{\partial E_3}{\partial t} + 2\alpha_2 E_2 \frac{\partial E_2}{\partial t} + 2\beta_2 E_3 \frac{\partial E_3}{\partial t} \end{aligned} \tag{6.154}$$

and the boundary conditions are given by (6.103) and (6.105).

Proposition 6.7.3 *If δ, a are sufficiently small, there exists a unique solution (E_2, E_3, H_2, H_3) of (6.103), (6.105) and (6.154), bounded and Lipschitz continuous together with its first derivatives in $D_a \times \mathbb{R}$. This solution is periodic in t with period $2\pi/\omega$, and depends continuously on the incident wave (in the uniform norm).*

For the proof, see [7] and the Appendix . As a Corollary, one can show that the perturbative solution is an actual approximation to the exact solution and derive informations concerning the second-harmonic wave.

The case of propagation parallel to the $y-$axis is similar.

6.7.9 Propagation along the optic axis.

For propagation parallel to the optic axis (the $z-$axis) the incident wave is given by eqs. (6.87) and (6.135), with the polarization angle θ defined as the angle between the $y-$axis and the electric field of the incident laser wave, like in §6.7.6. Eqs. (6.88), (6.89) and (6.152) yield

$$H_3 \equiv 0 \text{ everywhere} \quad , \qquad D_3 = \epsilon_o E_3 \equiv 0 \qquad \text{for } z < 0 \, , \, z > a$$

and

$$D_3 = \epsilon_{\|} E_3 + \alpha_2 (E_1^2 + E_2^2) + \beta_2 E_3^2 \equiv 0 \ \text{ for } \ 0 \leq z \leq a$$

where $|\boldsymbol{E}|^2 := E_1^2 + E_2^2$, so that the longitudinal electric field in the crystal is

$$E_3 = \frac{\epsilon_{\|}}{2\beta_2} \left[\sqrt{1 - 4\frac{\alpha_2 \beta_2}{\epsilon_{\|}^2} |\boldsymbol{E}|^2} - 1 \right] \equiv -2\frac{\alpha_2}{\epsilon_{\|}} |\boldsymbol{E}|^2 + O(\delta^2) \tag{6.155}$$

$\boldsymbol{E}$ and $\boldsymbol{H}$ are unknown vector functions of (z, t) of the form

$$\boldsymbol{E}(z,t) = E_1(z,t)\mathbf{c}_1 + E_2(z,t)\mathbf{c}_2 + E_3(z,t)\mathbf{c}_3 \quad , \quad \boldsymbol{H}(z,t) = H_1(z,t)\mathbf{c}_1 + H_2(z,t)\mathbf{c}_2$$

where E_3 is given by eq. (6.155) inside the crystal and vanishes outside. The basic unknowns are the fields E_1, E_2, H_1 and H_2 inside the slab, which satisfy the nonlinear Maxwell equations

$$\begin{aligned} &\epsilon_\perp \frac{\partial E_1}{\partial t} + 2\beta_1 \frac{\partial (E_2 E_3)}{\partial t} + 2\alpha_1 \frac{\partial (E_1 E_3)}{\partial t} = -\frac{\partial H_2}{\partial z} \ , \quad \mu_o \frac{\partial H_2}{\partial t} = -\frac{\partial E_1}{\partial z} \\ &\epsilon_\perp \frac{\partial E_2}{\partial t} + 2\alpha_1 \frac{\partial (E_2 E_3)}{\partial t} - 2\beta_1 \frac{\partial (E_1 E_3)}{\partial t} = \frac{\partial H_1}{\partial z} \ , \quad \mu_o \frac{\partial H_1}{\partial t} = \frac{\partial E_2}{\partial z} \end{aligned} \tag{6.156}$$

for $0 < z < a$, $t \in \mathbb{R}$ and the boundary conditions (6.143)

$$\begin{aligned}
&\sqrt{\epsilon_o/\mu_o} E_1(0,t) + H_2(0,t) = 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, sin\,\theta \\
&\sqrt{\epsilon_o/\mu_o} E_2(0,t) - H_1(0,t) = 2\sqrt{\epsilon_o/\mu_o}\; E_l\, cos(\omega t)\, cos\,\theta \\
&\sqrt{\epsilon_o/\mu_o} E_1(a,t) - H_2(a,t) = 0\ , \quad \sqrt{\epsilon_o/\mu_o} E_2(a,t) + H_1(a,t) = 0
\end{aligned}$$

for all t. Unfortunately, as already mentioned, an existence and uniqueness theory in the nonlinear case is not available here, because there is no way to reduce the nonlinear hyperbolic system (6.156) to characteristic form with a diagonally dominant matrix (see §6.7.6 and the Appendix). On the other hand, the linearized system uncouples into the two 2×2 systems

$$\begin{aligned}
\epsilon_\perp \frac{\partial E_1^o}{\partial t} &= -\frac{\partial H_2^o}{\partial z} \quad , \; \mu_o \frac{\partial H_2^o}{\partial t} = -\frac{\partial E_1^o}{\partial z} \\
\epsilon_\perp \frac{\partial E_2^o}{\partial t} &= \frac{\partial H_1^o}{\partial z} \ , \; \mu_o \frac{\partial H_1^o}{\partial t} = \frac{\partial E_2^o}{\partial z}
\end{aligned}$$

to which Cesari's theory does apply, and the linearized solution

$$E_1^o = E_o(z,t) sin\,\theta,\ H_2^o = -H_o(z,t) sin\,\theta\ , E_2^o = E_o(z,t)\, cos\,\theta,\ H_1^o = H_o(z,t)\, cos\,\theta$$

is the same as in §6.7.6. If we then write the formal perturbation expansions

$$\begin{aligned}
E_1 &= E_1^o + \delta E_1'(z,t) + O(\delta^2)\ , \quad H_2 = H_2^o + \delta H_2'(z,t) + O(\delta^2) \\
E_2 &= E_2^o + \delta E_2'(z,t) + O(\delta^2)\ , \quad H_1 = H_1^o + \delta H_1'(z,t) + O(\delta^2)
\end{aligned}$$

we find that here, by force of (6.154),

$$E_3 = O(\delta)$$

(for fixed E_l). The perturbations $E_1', ..., H_2'$ up to first order in δ then satisfy the homogeneous linear system

$$\begin{aligned}
\epsilon_\perp \frac{\partial E_1'}{\partial t} + \frac{\partial H_2'}{\partial z} &= 0\ , \quad \mu_o \frac{\partial H_2'}{\partial t} + \frac{\partial E_1'}{\partial z} = 0 \\
\epsilon_\perp \frac{\partial E_2'}{\partial t} - \frac{\partial H_1'}{\partial z} &= 0\ , \quad \mu_o \frac{\partial H_1'}{\partial t} - \frac{\partial E_2'}{\partial z} = 0
\end{aligned}$$

and the homogeneous boundary conditions (6.150). It follows that they are identically zero

$$E_1'(z,t) \equiv H_1'(z,t) \equiv E_2'(z,t) \equiv H_2'(z,t) \ \equiv 0$$

and so the solution in the slab coincides with the linearized solution up to terms of order δ^2:

$$E_1 = E_1^o + O(\delta^2)\ ,\ \ H_2 = H_2^o + O(\delta^2)$$
$$E_2 = E_2^o + O(\delta^2)\ ,\ \ H_1 = H_1^o + O(\delta^2)$$

Therefore the second-harmonic wave is (at most) of order δ^2, and hence is too small to be observed in practice. In other words, for propagation along the optic axis of a crystal of class 6-C6 (like $LiIO_3$) the perturbative solution predicts no observable second-harmonic generation.

Remark 8. The above process can be extended to include dispersion effects for the linear permittivity and the nonlinear polarizability: the boundary conditions on the slab walls remain exactly the same, since they depend only on the electromagnetic field in the non-dispersive medium outside the slab. However, the corrections to the second-harmonic wave due to dispersion turn out to be very small (of order $O(\delta^2)$, see §A.7 in the Appendix).

Remark 9. In the recent physics literature, nonlinear optics is often dealt with by means of nonlinear Schrödinger equations obtained by neglecting some terms in the nonlinear Maxwell equations. This approach works for any geometry and in any number of space variables, but involves an approximation for which no rigorous error estimates are available. In contrast, Cesari's method is exact and is based on a rigorous mathematical analysis and quantitative error estimates, but works only in one space variable. If extended to wave propagation problems in two or three space dimensions Cesari's approach would be no longer exact, as the impedance boundary conditions become approximate ones [32,35], but a uniqueness theorem can still be proven [21].

APPENDIX to section 7: THE NONLINEAR BOUNDARY VALUE PROBLEM

A.1 Characteristic form for hyperbolic systems. Consider the quasilinear system in two independent variables z, t

$$\mathbf{u}_z + \mathbb{A}(z,t,\boldsymbol{u})\boldsymbol{u}_t = \boldsymbol{w}(z,t,\boldsymbol{u}) \tag{6.157}$$

where $\mathbf{u} = \mathbf{u}(z,t) = (u_1, ..., u_m)$ is the unknown vector function, with values in Ω, a bounded domain in $\mathbb{R}^m$, and (z,t)$\in$D$_a \times \mathbb{R}, D_a = \{0 \leq z \leq a\}$. We denote here by $\mathbf{w} = (w_1, ..., w_m)$ a given vector function, bounded and continuous in $D_a \times \mathbb{R} \times \Omega$, and by $\mathbb{A}(z,t,u)$ an $m \times m$ matrix whose entries are bounded and continuous in $D_a \times \mathbb{R} \times \Omega$.

We suppose that the system (6.157) is hyperbolic in the $z-$variable, in the sense that the matrix $\mathbb{A}(z,t,\boldsymbol{u})$ has a full set of left eigenvectors [9] $\boldsymbol{h}_i = \boldsymbol{h}_i(z,t,\boldsymbol{u}) = (h_i^1, ..., h_i^m)$ in $\mathbb{R}^m$ corresponding to the eigenvalues $\rho_i(z,t,\boldsymbol{u})$:

$$\boldsymbol{h}_i \mathbb{A} = \rho_i\, \boldsymbol{h}_i \qquad (i = 1, ..., m) \tag{6.158}$$

at every point $(z,t,\boldsymbol{u}) \in D_a \times \mathbb{R} \times \Omega$ [2]. Let $\mathbb{S} = \mathbb{S}(z,t,\boldsymbol{u})$ denote the non-singular matrix whose rows are the m eigenvectors $\boldsymbol{h}_i$:

$$\mathbb{S}(z,t,\boldsymbol{u}) := \begin{pmatrix} h_1^1 & .... & h_1^m \\ .... & .... & .... \\ h_m^1 & .... & h_m^m \end{pmatrix} \tag{6.159}$$

In other words, the entries of $\mathbb{S}$ are $S_{ij} = h_i^j$. Eq. (6.158) can then be written as

$$\mathbb{S}\mathbb{A} = \mathbb{D}\mathbb{S}$$

where

$$\mathbb{D}(z,t,\boldsymbol{u}) := \begin{pmatrix} \rho_1 & .... & 0 \\ .... & .... & .... \\ 0 & .... & \rho_m \end{pmatrix}$$

is the $m \times m$ diagonal matrix with eigenvalues $\rho_1, ..., \rho_m$. Hence

$$\mathbb{A} = \mathbb{S}^{-1}\mathbb{D}\mathbb{S}$$

so that, denoting $\boldsymbol{W} = \mathbb{S}\boldsymbol{w}$, eq. (6.157) takes the characteristic form

$$\mathbb{S}(z,t,\boldsymbol{u})\boldsymbol{u}_z + \mathbb{D}\mathbb{S}(z,t,\boldsymbol{u})\,\boldsymbol{u}_t = \boldsymbol{W}(z,t,\boldsymbol{u})$$

In terms of components we have

$$\sum_{j=1}^{m} S_{ij}(z,t,\boldsymbol{u})\Big[\frac{\partial u_j}{\partial z} + \rho_i(z,t,\boldsymbol{u})\frac{\partial u_j}{\partial t}\Big] = W_i(z,t,\boldsymbol{u}) \quad (i = 1, ..., m) \tag{6.160}$$

[9] it can be proven [18] that, interchanging z and t, the symmetric hyperbolic systems introduced in Chapter 4 are also hyperbolic in the sense of this definition

where $\boldsymbol{u}$=$\boldsymbol{u}(z,t)$. In this form, the left-hand side of the $i-$th equation (6.160) is a linear combination of the derivatives of u_j along the $i-$th characteristic curve $t = t(z)$ satisfying

$$\frac{dt}{dz} = \rho_i(z,t,\boldsymbol{u}) \tag{6.161}$$

If $\mathbb{S} = \mathbb{I}$, i.e. $S_{ij} = \delta_{ij}$, the characteristic form reduces to the particular case of the diagonal form

$$\frac{\partial u_i}{\partial z} + \rho_i(z,t,\mathbf{u})\frac{\partial u_i}{\partial t} = W_i(z,t,\mathbf{u}), \quad (i = 1,...,m)$$

and the u_i 's are called Riemann invariants [2,18]. Note that here, since the space and time variables are interchanged, and z instead of t is taken as the hyperbolic variable, the eigenvalues

$$\rho_i(z,t,\boldsymbol{u}) = 1/s_i(z,t,\boldsymbol{u})$$

denote the inverse of the characteristic speeds s_i of Chapter 4.

The above discussion shows that a quasilinear hyperbolic system in two variables can always be reduced (locally) to characteristic form. The reduction, though, is not unique, since the matrix $\mathbb{S}$ can be multiplied by an arbitrary invertible matrix, which amounts to assuming as new unknown functions arbitrary one-to-one linear substitutions of the u_j's. We will revert to this important point later on.

In contrast, for hyperbolic systems in three or more independent variables there is a possibility of "loss of derivatives", related to the physical phenomenon of "focussing" [18], and the reduction to characteristic form is in general impossible.

Proposition 6.7.4 *The quasilinear hyperbolic system of m equations in $r+1$ independent variables*

$$\boldsymbol{u}_{x_o} + \sum_{k=1}^{r} \mathbb{A}_k(\boldsymbol{x},\boldsymbol{u})\,\boldsymbol{u}_{x_k} = \boldsymbol{w}(\boldsymbol{x},\boldsymbol{u}) \tag{6.162}$$

(where $\boldsymbol{u}$= $(u_1,...,u_m)$ =$\boldsymbol{u}(\boldsymbol{x})$, $\boldsymbol{x}$= $(x_o,...,x_r)$, and x_o is the hyperbolic variable) can be reduced to characteristic form

$$\sum_{j=1}^{m} S_{ij}(\boldsymbol{x},\boldsymbol{u})\Big[\frac{\partial u_j}{\partial x_o} + \sum_{k=1}^{r} \rho_{i,k}(\boldsymbol{x},\boldsymbol{u})\frac{\partial u_j}{\partial x_k}\Big] = W_i(\boldsymbol{x},\boldsymbol{u}) \quad (i = 1,...,m) \tag{6.163}$$

if and only if the r matrices $\mathbb{A}_k(\boldsymbol{x},\boldsymbol{u})$ commute.

Proof. By the definition of hyperbolicity, the matrices $\mathbb{A}_k$ have full sets of eigenvectors in in $\mathbb{R}^m$ for every $(\boldsymbol{x},\boldsymbol{u})$. If in addition $\mathbb{A}_k\mathbb{A}_j = \mathbb{A}_j\mathbb{A}_k$ for all $j,k = 1,...,r$, they can be diagonalized simultaneously, i.e. they have a full set of eigenvectors $\boldsymbol{h}_i = \boldsymbol{h}_i(\boldsymbol{x},\boldsymbol{u})$ in common. Let $\rho_{i,k}$ $(i = 1,... ..,m,\ k = 1,...,r)$ denote the $i-th$ eigenvalue of the matrix $\mathbb{A}_k$, and let $\mathbb{D}_k$ be the diagonal matrix with eigenvalues $\rho_{i,k}$. If $\mathbb{S}$ is the eigenvector matrix as in (6.159), then $\mathbb{S}\mathbb{A}_k = \mathbb{D}_k\mathbb{S}$ and from (6.162) we obtain

$$\mathbb{S}\boldsymbol{u}_{x_o} + \sum_{k=1}^{r} \mathbb{D}_k(\boldsymbol{x},\boldsymbol{u})\,\mathbb{S}(\boldsymbol{x},\boldsymbol{u})\boldsymbol{u}_{x_k} \equiv \mathbb{S}\boldsymbol{w}(\boldsymbol{x},\boldsymbol{u}) \tag{6.164}$$

which is the characteristic form (6.163) in vector notation. Conversely, if the system has the form (6.164) the matrices $\mathbb{A}_k(\boldsymbol{x},\boldsymbol{u}) := \mathbb{S}^{-1}\mathbb{D}_k\mathbb{S}$ clearly commute.

Corollary 6.7.5 *The Maxwell equations in more than one space variable cannot be reduced to characteristic form.*

Proof. The matrices $\mathbb{A}_k$ for the Maxwell equations do not commute (see eqs. (83), (84) of Chapter 4). The characteristic form implies absence of "focussing" and of "loss of derivatives", and for the Maxwell equations this is true only in one space dimension.

Remark 10. The m families of curves depending on $\boldsymbol{u}\in\Omega$

$$\frac{dx_k}{dx_o} = \rho_{i,k}(\boldsymbol{x},\boldsymbol{u}) \qquad (k = 1,...,r) \tag{6.165}$$

$(i = 1,...,m)$ are called the bicharacteristic rays of the hyperbolic system (6.162) (see eq. (4.122) and Exercise 15). Each equation (6.163) contains a linear combination of derivatives along the $i-th$ bicharacteristic ray. If $r = 1$, characteristics and bicharacteristics coincide.

A.2 Cesari's theorem. We consider from now on the hyperbolic system (6.160) in characteristic form with two independent variables $(z,t) \in D_a \times \mathbb{R}$. We will make the following additional assumptions.

H7. The eigenvectors $\boldsymbol{h}_1(z,t,\boldsymbol{u}), ..., \boldsymbol{h}_m(z,t,\boldsymbol{u})$, the eigenvalues $\rho_i(z,t,\boldsymbol{u})$ and the vector functions $\boldsymbol{w}(z,t,\boldsymbol{u})$ are bounded and uniformly Lipschitz continuous over all of $D_a \times \mathbb{R} \times \Omega$. Moreover,we assume that $det\, \mathbb{S}(z,t,\boldsymbol{u}) \geq M > 0$ for all $(z,t,\boldsymbol{u}) \in D_a \times \mathbb{R} \times \Omega$, and we denote by $\boldsymbol{\psi} = (\psi_1(t), ..., \psi_m(t))$ a bounded and uniformly Lipschitz continuous vector function for $t \in \mathbb{R}$:

$$|\psi_i(t)| \leq b \ , \ \ |\psi_i(t) - \psi_i(t')| \leq L|t - t'|$$

$(b > 0,\ L > 0,\ i = 1, ..., m)$.

These assumptions imply that the entries $S_{ij}(z,t,\boldsymbol{u})$, $s_{ij}(z,t,\boldsymbol{u})$ of the matrices $\mathbb{S}$ and $\mathbb{S}^{-1}$, respectively, are also bounded and uniformly Lipschitz continuous. Thus, there exists a constant $\Lambda > 0$ such that

$$|S_{ij}(z,t,\boldsymbol{u}) - S_{ij}(z',t',\boldsymbol{u}')| \leq \Lambda\Big[|z - z'|+|t - t'|+\|\boldsymbol{u} - \boldsymbol{u}'\|\Big]$$
$$|s_{ij}(z,t,\boldsymbol{u}) - s_{ij}(z',t',\boldsymbol{u}')| \leq \Lambda\Big[|z - z'|+|t - t'|+\|\boldsymbol{u} - \boldsymbol{u}'\|\Big]$$

over all of $D_a \times \mathbb{R} \times \Omega$ and for all $i,j = 1, ..., m$, where $\|.\|$ denotes some vector norm in $\mathbb{R}^m$.

We supply the system (6.160) with the following set of boundary conditions

$$\sum_{j=1}^{m} b_{ij} u_j(a_i,t) = \psi_i(t) \qquad\qquad (t \in \mathbb{R},\ i = 1, ..., m) \tag{6.166}$$

where $a_1, ..., a_m$ are m given numbers equal either to 0 or to a. Let $\mathbb{B}$ denotes the square matrix with entries b_{ij} (which in general might depend on t), and let $\tilde{\mathbb{S}}$, $\tilde{\mathbb{S}}^{-1}$, $\tilde{\mathbb{B}}$ denote the square matrices whose entries are defined by

$$\tilde{S}_{ij} := S_{ij} - \delta_{ij} \ , \ \tilde{s}_{ij} := s_{ij} - \delta_{ij},\ \ \tilde{b}_{ij} := b_{ij} - \delta_{ij} \tag{6.167}$$

$(i,j = 1, ..., m)$. Furthermore, let us define the nonnegative quantities

$$\sigma_o := sup \sum_{j=1}^{m} |\,\tilde{b}_{ij}|,$$
$$\sigma_1 := sup \sum_{j=1}^{m} |\tilde{S}_{ij}(z,t,\boldsymbol{u})|, \tag{6.168}$$
$$\sigma_2 := sup \sum_{j=1}^{m} |\tilde{s}_{ij}(z,t,\boldsymbol{u})|$$

where the supremums are taken over all relevant $(i, z, t, \boldsymbol{u})$.

We are now in a position to state Cesari's theorem in a simplified version tailored to our needs.

Theorem 6.7.6 *[4, 15]. Under assumptions **H7**, let the matrices $\mathbb{S}$, $\mathbb{B}$ have dominant main diagonals in the sense that*

$$(\sigma_o + \sigma_1)(1 + \sigma_2) < 1 \tag{6.169}$$

Then if a is sufficiently small there exists one and only one bounded vector function $\mathbf{u}(z, t)$, uniformly Lipschitz continuous with respect to (z, t), which satisfies (6.160) a.e. in $D_a \times \mathbb{R}$ and (6.166) for all $t \in \mathbb{R}$:

$$|u_i(z,t)| \le M \quad |u_i(z,t) - u_i(z',t')| \le Q\left[|z - z'| + |t - t'|\right]$$

$(M > 0,\ Q > 0,\ i = 1, ..., m)$. *This solution is unique and depends continuously on $|\psi_i(t)|$ (in the uniform norm). If all $\psi_i(t)$ are periodic with period $2\pi/\omega$, the solution is periodic with the same period.*

Remark 11. The characteristic form is essential for Theorem 6.7.6, since it guarantees that there is no "loss of derivatives". The differential problem can then be formulated as a fixed point problem for a set of nonlinear integral equations and analyzed by means of a suitable version of the well-known Banach contraction mapping principle.

Remark 12. The restriction on δ mentioned in the statements of Proposition 6.7.1, Proposition 6.7.2 and Proposition 6.7.3 is made explicit by the assumption (6.169) of dominant main diagonals of the matrices $\mathbb{S}$ and $\mathbb{B}$. This is the key assumption of Cesari's theory. Simple counterexamples show that such an assumption is essential.

It is possible to prove a regularity theorem showing that the solution $\mathrm{u}(z, t)$ is a classical solution under additional smoothness hypotheses, which are satisfied in the case of the "laser problem".

Theorem 6.7.7 *[2]. In addition to the assumptions of Theorem 6.7.6, suppose that the first partial derivatives of the matrix $\mathbb{A}(z, t, \boldsymbol{u})$, of $\boldsymbol{w}(z, t, \boldsymbol{u})$ and of $\psi_i(t)$ are bounded and uniformly Lipschitz continuous over $D_a \times \mathbb{R} \times \Omega$ and $\mathbb{R}$, respectively. Then the a.e. solution $\mathbf{u}(z, t)$ has Lipschitz continuous first derivatives $\partial\mathbf{u}/\partial z$, $\partial\mathbf{u}/\partial t$ in $D_a \times \mathbb{R}$ and is a classical solution of the boundary value problem.*

Remark 13. The supremum of the set of admissible slab widths a can be thought of as a lower estimate for the critical span a_s at which discontinuous solutions or shocks may occur (see Remark 7). Note that, in Theorems 6.7.6 and 6.7.7, we denote by z a generic space variable in the interval $[0, a]$.

A.3 Application of Cesari's theorem to the "laser problem". In order to apply Theorems 6.7.6 and 6.7.7 to the appropriate Maxwell equations in the crystal, which we write in the form of the general quasilinear system (6.157)

$$\boldsymbol{u}_z + \mathbb{A}(z, t, \boldsymbol{u})\, \boldsymbol{u}_t = \boldsymbol{w}(z, t, \boldsymbol{u})$$

we must first reduce this system to the characteristic form (6.163) in such a way that the key assumption (6.169) of dominant main diagonals of the matrices $S_{ij}(z, t, \boldsymbol{u})$ and b_{ij} be satisfied. Let δ be the small parameter defined previously, in eq. (6.132) or (6.153). We may then proceed as follows:

(i) Write the quantity $(\sigma_o + \sigma_1)(1 + \sigma_2)$ for $\delta = 0$ (linearized case) and verify that it is less than one.

(ii) Verify that for a (reasonably) small $\delta > 0$ the nonlinear corrections are small, so that the quantity $(\sigma_o + \sigma_1)(1 + \sigma_2)$ remains smaller than one in the nonlinear case.

However, if one tries a naive implementation of these two steps, according to the recipee of section A.1, one finds that the resulting matrices $\mathbb{S}$ and $\mathbb{B}$ are in general not diagonally dominant, even for $\delta = 0$. On the other hand, we have seen that the characteristic form is not unique. Hence one can try to look for a one-to-one linear substitution $\mathbb{L}$ of the dependent variables

$$\boldsymbol{U} = \mathbb{L}(\boldsymbol{u}) \quad \Leftrightarrow \quad \boldsymbol{u} = \mathbb{L}^{-\mathbf{1}}(\boldsymbol{U})$$

such that the transformed matrices $\mathbb{S}(z, t, \boldsymbol{U})$ and $\mathbb{B}$, corresponding to the new dependent variables $\boldsymbol{U}$, do satisfy (6.169).

This program is easy to carry out when one has to deal with 2×2 systems ($m = 2$). We have seen above that in many cases one has to solve a nonlinear system in two unknowns whose linearized version (for $\delta = 0$) has the form

$$\frac{\partial H}{\partial x} = \epsilon \frac{\partial E}{\partial t} \quad , \quad \frac{\partial E}{\partial x} = \mu_o \frac{\partial H}{\partial t} \tag{6.170}$$

where ϵ is either $\epsilon_\perp$ or $\epsilon_\parallel$ (see eqs. (6.94), (95), (6.97), (6.104), (6.112), (6.122), (6.127), (6.130)). The matrix $\mathbb{B}$ is then a 2×2 matrix of the form

$$\mathbb{B} = \begin{pmatrix} \sqrt{\frac{\epsilon_o}{\mu_o}} & -1 \\ \sqrt{\frac{\epsilon_o}{\mu_o}} & 1 \end{pmatrix}$$

(or can be splitted into two such matrices), corresponding to the boundary conditions

$$\sqrt{\frac{\epsilon_o}{\mu_o}}\, E(0,t) - H(0,t) = \Psi(t) \ , \quad \sqrt{\frac{\epsilon_o}{\mu_o}} E(a,t) + H(a,t) = 0 \tag{6.171}$$

where $\Psi(t)$ is a given function (see eqs. (6.103), (6.105), (6.113), (6.120), (6.128), (6.129), (6.131), (6.143)). Clearly, in order to check assumption (6.169) one must first homogenize the physical dimensions. This can be done by defining the variables [10]

$$\tau := ct \ , \quad u(z,\tau) := \sqrt{\frac{\epsilon}{\mu_o}}\, E(x,t) \ , \ \ v(z,\tau) := H(x,t)$$

where $c := (\epsilon\mu_o)^{-1/2}$. Eq. (6.170) becomes then

$$\frac{\partial u}{\partial \tau} = \frac{\partial v}{\partial x} \ , \quad \frac{\partial v}{\partial \tau} = \frac{\partial u}{\partial x} \tag{6.172}$$

and the matrix $\mathbb{B}$ changes into

$$\mathbb{B} = \begin{pmatrix} h & -1 \\ h & 1 \end{pmatrix} \quad \Rightarrow \quad \tilde{\mathbb{B}} = \begin{pmatrix} h-1 & -1 \\ h & 0 \end{pmatrix}$$

where the non-dimensional parameter

$$h := \sqrt{\epsilon_o/\epsilon}$$

satisfies $0 < h < 1$. This matrix $\mathbb{B}$ is not diagonally dominant; precisely, it is immediate to check that $\sigma_o = 2 - h > 1$.

Therefore new dependent variables

$$\boldsymbol{U} = \mathbb{L}(\boldsymbol{u}), \ \ \boldsymbol{U} = (U,V), \ \boldsymbol{u} = (u,v)$$

[10] τ has dimension length like z , u and v have the same dimension (of magnetic field)

are needed. We choose for $\mathbb{L}$ the linear one-to-one transformation

$$(6.173) \qquad \begin{aligned} U(x,\tau) &:= \frac{1}{2}\left[u(x,\tau)+v(x,\tau)\right] \equiv \frac{1}{2}\Big[\sqrt{\frac{\epsilon}{\mu_o}}E(x,t)+H(x,t)\Big] \\ V(x,\tau) &:= \frac{1}{2}\left[u(x,\tau)-v(x,\tau)\right] \equiv \frac{1}{2}\Big[\sqrt{\frac{\epsilon}{\mu_o}}E(x,t)-H(x,t)\Big] \end{aligned}$$

Substituting these equations in (6.172), we obtain the system in diagonal form

$$(6.174) \qquad \frac{\partial U}{\partial \tau}=\frac{\partial U}{\partial x} \quad \frac{\partial V}{\partial \tau}=-\frac{\partial V}{\partial x}$$

whose general solution (in any convex set of $\mathbb{R}^2$) is given by

$$U=\Phi(y+\tau) \quad , \quad V=\Psi(y-\tau)$$

with Φ and Ψ arbitrary C^1 functions (U, V are the linearized Riemann invariants in the crystal [2, 18]). Since the transformed system (6.174) is in diagonal form, the corresponding matrices $\mathbb{S}$, $\mathbb{S}^{-1}$ will coincide with the identity matrix and therefore

$$\sigma_1=\sigma_2=0$$

Similarly, taking U and V as dependent variables the new matrix $\mathbb{B}$ reads

$$\mathbb{B}=\begin{pmatrix} \frac{1}{2}(1+h) & \frac{1}{2}(h-1) \\ \frac{1}{2}(h-1) & \frac{1}{2}(1+h) \end{pmatrix} \quad \Rightarrow \quad \tilde{\mathbb{B}}=\begin{pmatrix} \frac{1}{2}(h-1) & \frac{1}{2}(h-1) \\ \frac{1}{2}(h-1) & \frac{1}{2}(h-1) \end{pmatrix}$$

Since $h>0$, $\mathbb{B}$ has dominant main diagonal and

$$\sigma_o=1-h\ <1$$

Summarizing, for a linear 2×2 system, if the linearized Riemann invariants U, V are taken as dependent variables, the condition of dominant main diagonals (6.169) is satisfied with

$$(\sigma_o+\sigma_1)(1+\sigma_2)=\sigma_o\equiv 1-h \quad , \quad 0<h<\ 1$$

For the nonlinear system corresponding to (6.170)

$$(6.175) \qquad \frac{\partial H}{\partial x}=(\epsilon-2\alpha E)\,\frac{\partial E}{\partial t} \quad , \quad \frac{\partial E}{\partial x}=\mu_o\frac{\partial H}{\partial t}$$

we have $\delta = |\alpha| > 0$ and it is reasonable to foresee that, using the same dependent variables (U, V), the matrices $\mathbb{S}$, $\mathbb{S}^{-1}$ ought to be small perturbations of the identity matrix, so that

$$\sigma_i = O(\delta) \qquad (i = 1, 2)$$

and, by taking δ small enough, one should have

$$(\sigma_o + \sigma_1)(1 + \sigma_2) \equiv 1 - h + O(\delta) < 1$$

Thus assumption (6.169) would be satisfied, and Theorems 6.7.6, 6.7.7 could be applied.

This program actually does work out as sketched, with admissible values of a, δ agreeing with experiments, except in the case of propagation along the optic axis.

The situation is less simple if $m = 4$, as for propagation along the $x-$axis of a $6-$C6 crystal. In this case we have obtained the 4×2 nonlinear system (6.154) whose linearized version ($\delta = 0$) splits into the two uncoupled 2×2 systems

$$-\frac{\partial E_2}{\partial x} = \mu_o \frac{\partial H_3}{\partial t} \ , \quad -\frac{\partial H_3}{\partial x} = \epsilon_\perp \frac{\partial E_2}{\partial t}$$
$$\frac{\partial E_3}{\partial x} = \mu_o \frac{\partial H_2}{\partial t} \ , \quad \frac{\partial H_2}{\partial x} = \epsilon_\parallel \frac{\partial E_3}{\partial t}$$

Let us take here as new dependent variables the quantities $\boldsymbol{U} = \mathbb{L}(\boldsymbol{u})$ defined by the linear one-to-one substitution

(6.176)

$$U_1(x,\tau) := \frac{1}{2}\Big[\sqrt{\frac{\epsilon_\perp}{\mu_o}} E_2(x,t) - H_3(x,t)\ \Big] \ , \ U_2(x,\tau) := \frac{1}{2}\Big[\sqrt{\frac{\epsilon_\perp}{\mu_o}} E_2(x,t) + H_3(x,t)\Big]$$
$$U_3(x,\tau) := \frac{1}{2}\Big[\sqrt{\frac{\epsilon_\perp}{\mu_o}} E_3(x,t) - H_2(x,t)\ \Big] \ , \ U_4(x,\tau) := \frac{1}{2}\Big[\sqrt{\frac{\epsilon_\perp}{\mu_o}} E_3(x,t) + H_4(x,t)\Big]$$

where $\tau := c_\perp t$. The constant matrix $\tilde{\mathbb{B}}$ has then the form

$$(\,\tilde{b}_{ij}) = \begin{pmatrix} \frac{1}{2}(h-1) & \frac{1}{2}(h-1) & 0 & 0 \\ \frac{1}{2}(h-1) & \frac{1}{2}(h-1) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(h-1) & \frac{1}{2}(h-1) \\ 0 & 0 & \frac{1}{2}(h-1) & \frac{1}{2}(h-1) \end{pmatrix}$$

where

$$h := \sqrt{\epsilon_o/\epsilon_\perp} \ , \ \ 0 < h < 1$$

is the inverse of the ordinary refractive index $n_{or} > 1$, so that

$$\sigma_o = 1 - h \; < 1$$

as before. The matrix $\tilde{\mathbb{S}}(z, t, \boldsymbol{U})$ for $\delta = 0$ turns out to be given by

$$(\tilde{S}_{ij}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \zeta \\ 0 & 0 & \zeta & 0 \end{pmatrix}$$

where

$$\zeta := \frac{1-k}{1+k} \quad , \quad k := \frac{\epsilon_\parallel}{\epsilon_\perp}$$

After some easy calculations we find

$$\sigma_1 = |\zeta| \quad , \quad \sigma_2 = \frac{|\zeta|}{1 - |\zeta|}$$

Summarizing, for propagation along the $x-$axis of a 6C−6 crystal in the linearized case ($\delta = 0$) we have

$$(\sigma_o + \sigma_1)(1 + \sigma_2) = \frac{1 - h \, + |\zeta|}{1 - |\zeta|}$$

As $|\zeta| \ll 1$ [41], this quantity turns out to be less than 1, as requested. For example, for a $LiIO_3$ crystal we have

$$\zeta \cong 0.041, \quad h \cong 0.52, \quad k \cong 0.92, \quad 1 - k^2 \cong 0.15$$

whence

$$(\sigma_o + \sigma_1)(1 + \sigma_2) \; \cong 0.54$$

In the nonlinear case ($\delta > 0$), using the same variables $(U_1, ..., U_4)$, σ_1 and σ_2 ought to be of the form

$$\sigma_1 = |\zeta| \, + O(\delta) \quad , \quad \sigma_2 = \frac{|\zeta|}{1 - |\zeta|} \, + O(\delta)$$

so that

$$(\sigma_o + \sigma_1)(1 + \sigma_2) \equiv \frac{1 - h \, + |\zeta|}{1 - |\zeta|} + O(\delta) < 1 \tag{6.177}$$

provided δ is small enough. Then assumption (6.169) would be satisfied, and Theorems 6.7.6, 6.7.7 could be applied. Again, it can be shown that, in this and similar cases, things do work out as expected with admissible values of a, δ in agreement with experiments.

The $O(\delta)$ term in (6.177), however, contains divisors of the type $(1-k^2)$ which vanish for propagation along the optic axis, for then k must be replaced by 1. In this case the entire process breaks down, and Cesari's theory cannot be applied. A similar breakdown occurs for crystals of class 32−D3, like quartz. As mentioned in the text, for propagation along the optic axis of a uniaxial crystal there seems to be no way to reduce the nonlinear system to characteristic form in such a way that the matrix $\mathbb{S}$ has dominant main diagonal (Exercise 16 and Exercise 17). This obstruction, which arises for all uniaxial crystals, appears to be related to the fact that the nonlinear hyperbolic system (6.144), which is not symmetric, has characteristic curves whose multiplicity is not constant in the entire domain $D_a \times \mathbb{R} \times \Omega$ and depends on the values of the solution (E_1, E_2) and of the parameter δ. Indeed, the linearized system (6.145) uncouples into two (strictly hyperbolic) systems in two unknowns (E_1, H_2), (E_2, H_1), respectively, each of which has characteristic speeds $\pm c_\perp$ and characteristic curves

$$z \pm c_\perp t = \mathit{constant}$$

with $c_\perp = (\epsilon_\perp \mu_o)^{-\frac{1}{2}}$. The characteristic curves of the nonlinear system (6.144) then "bifurcate" from these multiple characteristics of the linearized system, and degenerate if $\alpha E_1(z,t) = \alpha E_2(z,t) = 0$ (Exercise 18).

Remark 14. One might introduce the deviations $\tilde{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{u}_o$, where $\boldsymbol{u}_o$ denotes the linearized solution [11], and modify the step (i) above by considering the matrix $\mathbb{S}(z, t, \boldsymbol{u}_o)$ (for $\delta > 0$) instead of the linearized matrix (with $\delta = 0$). This approach yields better estimates for the admissible slab width a, but the above qualitative conclusions remain unaltered.

A.4 Counterexamples. The boundedness assumption and the condition of diagonal dominance (6.169) in Theorems 6.7.6 and 6.7.7 are essential for existence and uniqueness, as the following simple example shows.

Consider the linear 2×2 system in diagonal form (6.174)

$$\frac{\partial U}{\partial \tau} = \frac{\partial U}{\partial x} \;, \quad \frac{\partial V}{\partial \tau} = -\frac{\partial V}{\partial x} \qquad (0 < x < a,\ \tau \in \mathbb{R})$$

[11] Ω will be then, typically, a functional ball centered on $\boldsymbol{u}_o$

together with the impedance boundary conditions at $x = 0, a$

$$\frac{1}{2}(h+1)U(0,\tau) + \frac{1}{2}(h-1)V(0,\tau) = \psi(\tau)$$
$$\frac{1}{2}(h-1)U(a,\tau) + \frac{1}{2}(h+1)V(a,\tau) = 0 \qquad (\tau \in \mathbb{R})$$

corresponding to (6.171). Setting

$$r := \frac{1-h}{1+h}$$

the boundary conditions can be written as

$$U(0,\tau) = rV(0,\tau) + 2\psi(\tau) \ , \quad V(a,\tau) = rU(a,\tau) \tag{6.178}$$

Suppose here that $0 \leq h < 1$, so that $0 < r \leq 1$. As we already know, the general solution of the differential system in any convex set of $\mathbb{R}^2$ is of the form

$$U = \Phi(\tau + x) \quad , V = \Psi(\tau - x) \tag{6.179}$$

so that from the boundary conditions (6.178) we obtain

$$\Phi(\tau) = r\Psi(\tau) + 2\psi(\tau) \ , \quad \Psi(\tau) = r\Phi(\tau + 2a) \qquad (\tau \in \mathbb{R})$$

and eliminating $\Psi(\tau)$ yields the linear functional equation for $\Phi(\tau)$

$$\Phi(\tau) = r^2\Phi(\tau + 2a) + 2\psi(\tau) \qquad (\tau \in \mathbb{R}) \tag{6.180}$$

Letting

$$\Phi(\tau) = r^{-\tau/a}P(\tau) \ \Rightarrow \ \Psi(\tau) = r^{-1}r^{-\tau/a}P(\tau + 2a)$$

eq. (6.180) can be reformulated as the linear difference equation for $P(\tau)$

$$P(\tau + 2a) - P(\tau) = -2r^{\tau/a}\psi(\tau) \quad (\tau \in \mathbb{R}) \tag{6.181}$$

(cfr. the Appendix to section 10 in Chapter 4). Solving the differential problem for U, V is therefore reduced to finding solutions $P(\tau) \in C^1(\mathbb{R})$ of the finite difference equation (6.181) for $\psi(\tau) \in C^1(\mathbb{R})$. (If $\psi(\tau)$ (hence $P(\tau)$) is only continuous or Lipschitz continuous, one obtains a weak solution or an a.e. solution of the boundary value problem, respectively.)

Case 1: Let $\psi(\tau) \equiv 0$ and $0 < h < 1$, so that the matrix $\mathbb{B}$ has dominant main diagonal, and eq. (6.181) becomes $P(\tau+2a) = P(\tau)$. Therefore we can take for $P(\tau)$ any periodic $C^1(\mathbb{R})-$function of period $2a$. As $\mathbb{S}$ is the identity matrix, all assumptions of Cesari's theorem are satisfied, but the previous equations show that there are infinitely many solutions of the homogeneous boundary value problem, given by

$$U(x,\tau) = r^{-(x+\tau)/a}P(\tau+x) \quad , \quad V(x,\tau) = r^{-1}r^{(x-\tau)/a}P(\tau-x) \tag{6.182}$$

By assumption, the reflection coefficient satisfies $0 < r < 1$, and so $U(x,\tau)$ and $V(x,\tau)$ are unbounded as $\tau \to +\infty$. Thus uniqueness fails if the boundedness assumption is dropped.

Case 2 : Let $h = 0$, so that $r = 1$ and the matrix $\mathbb{B}$ is not diagonally dominant. The previous equations show that $\Phi(\tau) = P(\tau)\,,\Psi(\tau) = P(\tau+2a)$ where $P(\tau)$ must be a C^1 solution of the difference equation

$$P(\tau+2a) - P(\tau) = -2\psi(\tau) \qquad (\tau \in \mathbb{R}) \tag{6.183}$$

Let us distinguish two subcases.

(*i*) If $\psi(\tau) \equiv 0$, eq. (6.183) implies that $P(\tau)$ is an arbitrary $2a-$periodic function of class $C^1(\mathbb{R})$ as in case 1, and there are infinitely many bounded solutions of the boundary value problem, given by eq. (6.182) with $r = 1$

$$U(x,\tau) = P(\tau+x) \quad , \quad V(x,\tau) = P(\tau-x)$$

and periodic in τ (and x) with period $2a$. In particular, the trigonometric functions

$$\begin{aligned} U(x,\tau) &= A_n cos(n\pi(\tau+x)/a) + B_n sin(n\pi(\tau+x)/a) \\ V(x,\tau) &= A_n cos(n\pi(\tau-x)/a) + B_n sin(n\pi(\tau-x)/a) \qquad (n = 0,1,...) \end{aligned}$$

corresponding to the eigenfunctions of a vibrating string of length a with fixed or free ends. Thus uniqueness fails if (6.169) is dropped.

(*ii*) If $\psi(\tau) \in C^1(\mathbb{R})$ is periodic with period $2a$ and not identically zero, eq. (6.183) implies that $P(\tau)$ cannot be periodic of period $2a$, and must satisfy

$$P(\tau+2na) = P(\tau) - 2\sum_{k=1}^{n-1}\psi(\tau+2ka) \equiv P(\tau) - 2(n-1)\psi(\tau)$$

for any integer n and all $\tau \in \mathbb{R}$. Take $\tau = \tau_o$ such that $\psi(\tau_o) \neq 0$. Then clearly $P(\tau_o + 2na)$ diverges as n goes to $+\infty$, U and V cannot be bounded, and the boundary value problem has no bounded or periodic solution (the problem is at resonance). Thus existence may fail if (6.169) is dropped.

Remark15. If $0 < r < 1$ the unique bounded solution of the difference equation (6.180) for any bounded source term $\psi(\tau)$ is given by the uniformly and absolutely convergent series

$$\Phi(\tau) = -2\sum_{n=0}^{\infty} r^{2n-2}\psi(\tau - 2a - 2na)$$

(see Exercise 16 of Chapter 4). It follows that

$$\Psi(\tau) = r\Phi(\tau + 2a) = -2r\sum_{n=0}^{\infty} r^{2n-2}\psi(\tau - 2na)$$

and the solution $U = \Phi(\tau + x)$, $V = \Psi(\tau - x)$ of the linear boundary value problem is

$$U = -2\sum_{n=0}^{\infty} r^{2n-2}\psi(\tau + x - 2a - 2na)\ , \quad V = -2r\sum_{n=0}^{\infty} r^{2n-2}\psi(\tau - x - 2na)$$

Continuous dependence of U, V upon ψ (in the uniform norm) follows easily from these formulae. Clearly, if $\psi(\tau)$ is periodic, U and V are periodic with respect to τ and x, with the same period. Moreover U and V have the same regularity as ψ (there is no "loss of derivatives") and in particular $\psi(\tau)$ must be of class $C^1(\mathbb{R})$ in order to obtain solutions U, V of class $C^1(\mathbb{R}^2)$.

Remark 16. If in the above example we set

$$U(x,\tau) := \frac{1}{2}\left[u(x,\tau) + v(x,\tau)\right]$$
$$V(x,\tau) := \frac{1}{2}\left[u(x,\tau) - v(x,\tau)\right]$$

then, for $\psi \equiv 0$, u and v satisfy the 2×2 system

$$u_\tau = v_x \quad , \ v_\tau = u_x \tag{6.184}$$

with the homogeneous impedance boundary conditions

$$hu(0,\tau) + v(0,\tau) = 0\ , \quad hu(a,\tau) - v(a,\tau) = 0 \tag{6.185}$$

Therefore u and v, if twice differentiable, are solutions of the homogeneous vibrating string equation for $0 < x < a$

$$u_{\tau\tau} = u_{xx} \quad , \quad v_{\tau\tau} = v_{xx}$$

with the homogeneous boundary conditions (6.185) at the ends $x = 0, a$ (and with $\tau = ct$).

If $h \neq 0$, the $L^2(0,a)-$norms $\|u\|$ and $\|v\|$ may diverge either as $\tau \to +\infty$ (if $h > 0$) or as $\tau \to -\infty$ (if $h < 0$), so that the string may undergo unbounded oscillations for $\tau \in \mathbb{R}$ in the absence of forcing terms (see eq. (6.182) and Exercise 19).

If $h = 0$ the norms $\|u\|$ and $\|v\|$ are time-independent, and eqs. (6.185) reduce to the boundary conditions $u_x = v = 0$ for $x = 0$ and $x = a$, giving rise to the well-known bounded eigenoscillations of a string of length a with fixed ends for v

$$v(x,\tau) = \sum_{n=1}^{\infty} (A_n cos\frac{n\pi\tau}{a} + B_n sin\frac{n\pi\tau}{a})\, sin\frac{n\pi x}{a}$$

and with free ends for u

$$u(x,\tau) = A_o - \sum_{n=1}^{\infty} (A_n sin\frac{n\pi\tau}{a} - B_n cos\frac{n\pi\tau}{a})\, cos\frac{n\pi x}{a}$$

where A_n and B_n $(n = 0, 1, ...)$ are arbitrary constants [2,18].

This remark may be of help in understanding the role of the boundedness and diagonal dominance assumptions. Such assumptions can of course be dispensed with in the case of an initial value problem, i.e. if U, V are both assigned at $x = 0$ (or at $x = a$).

A.5 The nonlinear BVP as limit of a sequence of linear BVP's. Consider the linear boundary value problem for the unknown vector function $\boldsymbol{u}(z,t)$

$$\sum_{j=1}^{m} S_{ij}(z,t,\boldsymbol{v}(z,t))\Big[\frac{\partial u}{\partial z} + \rho_i(z,t,\boldsymbol{v}(z,t))\frac{\partial u_j}{\partial t}\Big] = W_i(z,t,\boldsymbol{v}(z,t)) \tag{6.186}$$

$$\sum_{j=1}^{m} b_{ij} u_j(a_i,t) = \psi_i(t) \qquad (0 \leq a_i \leq a \;\; i = 1, ..., m) \tag{6.187}$$

where $\boldsymbol{v}(z, t)$ is an arbitrarily auxiliary vector function satisfying the same assumptions as $\boldsymbol{u}$. This differential problem can be formulated as a system of integral equations and the solution $\boldsymbol{u}$ can be envisaged as the fixed point of a linear integral transformation

$$\boldsymbol{U} = \mathbb{T}\boldsymbol{u}$$

where the map $\mathbb{T} = \mathbb{T}_v$ depends on $\boldsymbol{v}$. If $a > 0$ is small enough, $\mathbb{T}$ turns out to be a contraction in a suitable Banach space of functions with the uniform topology, so that there is a unique fixed point

$$\overline{\boldsymbol{u}} = \mathbb{T}_{\boldsymbol{v}}\, \overline{\mathbf{u}} \tag{6.188}$$

which depends on $\boldsymbol{v}$ and can be obtained as limit of an iterative sequence. Since the fixpoint $\overline{\boldsymbol{u}}$ of $\mathbb{T}_{\boldsymbol{v}}$ is easily seen to be differentiable, $\overline{\boldsymbol{u}}$ is the unique solution to (6.186), (6.187).

This process is standard in the case of a Cauchy problem, i.e. if all $a_i = 0\,(j = 1, ..., m)$.

We add here just a few more details, referring the reader to the quoted references for a complete discussion. The $i-$th equation (6.186) can be written in the form

$$\sum_{j=1}^{m} S_{ij}\, \frac{du_j}{dz_i} = W_i \tag{6.189}$$

where du_j/dz_i denotes the derivative of u_j along the $i-$th characteristic curve $t = t(z)$, satisfying the equation

$$\frac{dt}{dz} = \rho_i \qquad (i = 1, ..., m)$$

(see eq. (6.161)). Since $\rho_i = \rho_i(z, t, \boldsymbol{v}(z, t))$, the characteristic curves depend on $\boldsymbol{v}$. Under the assumptions of Theorem 6.7.7, S_{ij} has bounded and continuous first partial derivatives, and we may write (6.189) in the form

$$\frac{d}{dz_i} \sum_{j=1}^{m} S_{ij}\, u_j \;=\; \sum_{j=1}^{m} \frac{dS_{ij}}{dz_i} u_j + W_i$$

Integrating over the $i-$th characteristic curve from 0 to z yields

$$\sum_{j=1}^{m} S_{ij} u_j(z, t) = \sum_{j=1}^{m} \big[S_{ij}\, u_j\big|_{z=0} + \int_0^z (\sum_{j=1}^{m} \frac{dS_{ij}}{dz_i} u_j \, + W_i\,)\, dz$$

Multiplying by the entries $s_{ki}(z,t)$ of the inverse matrix $\mathbb{S}^{-1}$ and summing over i yields

$$u_k(z,t) = \sum_{i,j=1}^{m} s_{ki}(z,t)\big[S_{ij}\,u_j\big]_{z=0} \tag{6.190}$$
$$+ \sum_{i,j=1}^{m} s_{ki}(z,t)\int_0^z (\frac{dS_{ij}}{dz_i}u_j + W_i)\,dz$$

where $k = 1, ..., m$, and s_{ki}, S_{ij} depend also on $\boldsymbol{v}$. For the Cauchy problem, $u_j\big|_{z=0}$ are given data and the right-hand side of (6.190) defines the integral transformation $\mathbb{T}_{\boldsymbol{v}}$.

In the case of the boundary value problem the definition of $\mathbb{T}_{\boldsymbol{v}}$ must be modified by integrating from a_i to z in the right-hand-side of (6.190) and by replacing $\big[S_{ij}\,u_j\big]_{z=0}$ by

$$\big[S_{ij}\,u_j - \,b_{ij}u_j\big]_{z=a_i} + \psi_i(\tau)\delta_{ij}$$

where $\tau = \tau(t,z,a_i)$ is the time needed to go from z to a_i along the $i-$th characteristic curve. In this way one has $\tau = t$ when $z = a_i$, so that

$$u_k(a_i,t) = \sum_{i,j=1}^{m} s_{ki}(a_i,t)\Big[S_{ij}\,(a_i,t)u_j(a_i,t) - \,b_{ij}u_j(a_i,t) + \psi_i(t)\delta_{ij}\Big]$$
$$= u_k(a_i,t) - \sum_{i=1}^{m} s_{ki}(a_i,t)\Big[\sum_{j=1}^{m} b_{ij}u_j(a_i,t) - \psi_i(t)\Big]$$

It follows that

$$\sum_{i=1}^{m} s_{ki}(a_i,t)\Big[\sum_{j=1}^{m} b_{ij}u_j(a_i,t) - \psi_i(t)\Big] = 0 \quad (k = 1, ..., m)$$

and, since the matrix $\mathbb{S}^{-1}$ is invertible,

$$\sum_{j=1}^{m} b_{ij}u_j(a_i,t) - \psi_i(t) \;= 0 \quad (i = 1, ..., m)$$

If

$$\overline{\boldsymbol{u}} = \mathbb{T}_{\boldsymbol{v}}\overline{\boldsymbol{u}}$$

is a fixed point of $\mathbb{T}_{\boldsymbol{v}}$, we have just shown that $\overline{\boldsymbol{u}}$ satisfies the BC's (6.187) for all $\boldsymbol{v}$. Moreover, it is easy to see that $\overline{\boldsymbol{u}}$ is differentiable. The fixpoint equation for $\overline{\boldsymbol{u}} = (\overline{u}_1, ..., \overline{u}_m)$ reduces then to

$$\sum_{i,j=1}^{m} s_{ki}(z,t) \int_0^z (W_i - S_{ij}\frac{d\overline{u}_j}{dz_i})\, dz = 0$$

so that $\overline{\boldsymbol{u}}$ satisfies also eq. (6.189), that is to say the differential system (6.186). On the other hand the fixed point $\overline{\boldsymbol{u}} = \mathbb{T}_{\boldsymbol{v}}\overline{\boldsymbol{u}}$ exists and is unique since $\mathbb{T}_{\boldsymbol{v}}$ is a contraction, at least if a is small enough. Consider for $n = 0, 1, ...$ the sequence

$$\boldsymbol{u}^{(n)} = (u_1^{(n)}, ..., u_m^{(n)}) \;=\; \boldsymbol{u}^{(n)}(z,t)$$

where $\boldsymbol{u}^{(0)} \in \Omega$ is chosen arbitrarily and each $\boldsymbol{u}^{(n)}$ is the solution of the sequence of linear boundary value problems

$$(6.191)\quad \sum_{j=1}^{m} S_{ij}(z,t,\boldsymbol{u}^{(n-1)})\Big[\frac{\partial u_j^{(n)}}{\partial z} + \rho_i(z,t,\boldsymbol{u}^{(n-1)})\frac{\partial u_j^{(n)}}{\partial t}\Big] = W_i(z,t,\boldsymbol{u}^{(n-1)})$$

$$(6.192)\qquad \sum_{j=1}^{m} b_{ij}u_j^{(n)}(a_i,t) = \psi_i(t) \qquad (i = 1, ..., m)$$

for $(z,t) \in D_a \times \mathbb{R}$, $n = 1, 2, ...$.

Theorem 6.7.8 *Under the assumptions of Theorem 6.7.6, the sequence* $\boldsymbol{u}^{(n)}$ *converges uniformy on* $D_a\times\mathbb{R}$ *to the solution* $\hat{\boldsymbol{u}}$ *of the nonlinear BVP (6.160), (6.166)*:

$$\hat{\boldsymbol{u}}(z,t) \;=\; \lim_{n\to\infty} \boldsymbol{u}^{(n)}(z,t) \qquad \forall (z,t) \in D_a \times \mathbb{R}$$

Proof. The fixed point $\overline{\boldsymbol{u}} = \mathbb{T}_{\boldsymbol{v}}\,\overline{\mathrm{u}}$ depends on $\boldsymbol{v}$ and the map $\overline{\boldsymbol{u}} = \overline{\boldsymbol{u}}\big[\boldsymbol{v}\big]$ turns out to be a contraction in the uniform norm if a is small enough [4]. Hence there exists a unique $\boldsymbol{v}$= $\hat{\boldsymbol{u}}$ such that

$$\hat{\boldsymbol{u}} = \mathbb{T}_{\hat{\boldsymbol{u}}}\; \hat{\boldsymbol{u}}$$

This fixed point $\hat{\boldsymbol{u}}$ is the solution of the nonlinear BVP (6.160), (166), and by the Banach contraction mapping theorem, it can be obtained as the limit of the iteration sequence

$$\boldsymbol{u}^{(n)} = \mathbb{T}_{\boldsymbol{u}^{(n-1)}}\boldsymbol{u}^{(n-1)} \qquad n = 0, 1, 2, ...$$

which is immediately seen to satisfy (6.191), (6.192).

A.6 Comparison of the iterative and perturbative solutions.It is interesting to compare the convergent iterative method of section **A.5**, applied to the laser problem, with the (non-convergent) perturbative method at order $O(\delta)$, used in sections 6.7.4 and 6.7.9 . We take for simplicity the case of a 2×2 system like (6.175), with x replaced by z,

$$\frac{\partial H}{\partial z} = (\epsilon \; - 2\alpha E\,)\, \frac{\partial E}{\partial t}, \qquad \frac{\partial E}{\partial z} \; = \mu_o \frac{\partial H}{\partial t} \tag{6.193}$$

and with the boundary conditions (6.171)

$$\sqrt{\frac{\epsilon_o}{\mu_o}}\, E(0,t) - H(0,t) = \Psi(t) \;\;, \;\; \sqrt{\frac{\epsilon_o}{\mu_o}}\, E(a,t) + H(a,t) = 0 \tag{6.194}$$

We have then $m = 2, \boldsymbol{u}^{(n)} = (E^{(n)},\, H^{(n)})$, and $\delta = \alpha$. Let us assume $\boldsymbol{u}^{(0)} = 0$, i.e.

$$E^{(0)} \equiv \; H^{(0)} \equiv 0$$

as initial vector for the iterative scheme. Then the first iterate $\boldsymbol{u}^{(1)} = (E^{(1)},\, H^{(1)})$ satisfies the linear system

$$\frac{\partial H^{(1)}}{\partial z} = \epsilon \frac{\partial E^{(1)}}{\partial t} \;\;, \quad \frac{\partial E^{(1)}}{\partial z} \; = \; \mu_o \frac{\partial H^{(1)}}{\partial t}$$

and the boundary conditions corresponding to (6.194)

$$\sqrt{\frac{\epsilon_o}{\mu_o}}\, E^{(1)}(0,t) - H^{(1)}(0,t) = \Psi(t) \;\;, \;\; \sqrt{\frac{\epsilon_o}{\mu_o}}\, E^{(1)}(a,t) + H^{(1)}(a,t) = 0$$

and therefore coincides with the linearized solution. The second iterate $\boldsymbol{u}^{(2)} = (E^{(2)}, H^{(2)})$ satisfies the linear system

$$\frac{\partial H^{(2)}}{\partial z} = (\epsilon \; - 2\alpha E^{(1)}) \frac{\partial E^{(2)}}{\partial t} \;\;, \quad \frac{\partial E^{(2)}}{\partial z} \; = \mu_o \frac{\partial H^{(2)}}{\partial t} \tag{6.195}$$

together with the homogeneous boundary conditions corresponding to (6.171).

The perturbative solution up to order $O(\alpha)$ is obtained by setting

$$E = E^{(1)} + \tilde{E} \;\;, \quad H = H^{(1)} + \; \tilde{H} \;\;; \quad \tilde{E} \; = \alpha E' \,, \; \tilde{H} = \alpha H'$$

substitute in (6.193), (6.194) and cancel terms of order $O(\alpha^2)$. As a result we obtain the system for the deviations $(\tilde{E}, \tilde{H})$ in the perturbative solution

$$\frac{\partial \tilde{H}}{\partial z} - \epsilon \frac{\partial \tilde{E}}{\partial t} = -2\alpha E^{(1)} \frac{\partial E^{(1)}}{\partial t} \quad , \quad \frac{\partial \tilde{E}}{\partial z} = \mu_o \frac{\partial \tilde{H}}{\partial t} \tag{6.196}$$

In this system, the contribution due to the first step (linearized solution) appears solely as a driving term. On the other hand, if we make the same position in (6.195)

$$E^{(2)} = E^{(1)} + \tilde{E} \ , \quad H^{(2)} = H^{(1)} + \tilde{H}$$

we obtain the system for the second-iterate deviation

$$\frac{\partial \tilde{H}}{\partial z} - \epsilon \frac{\partial \tilde{E}}{\partial t} + 2\alpha E^{(1)} \frac{\partial \tilde{E}}{\partial t} = -2\alpha E^{(1)} \frac{\partial E^{(1)}}{\partial t}, \quad \frac{\partial \tilde{E}}{\partial z} = \mu_o \frac{\partial \tilde{H}}{\partial t} \tag{6.197}$$

which differs from (6.196) because of the presence of the additional term

$$2\alpha E^{(1)} \frac{\partial \tilde{E}}{\partial t}$$

in the first equation (6.197). This terms modifies the characteristic curves for $n = 2$ with respect to the linearized characteristic curves (for $n = 1$) :in the iterative scheme the characteristic curves are updated at each step. In contrast, the perturbative method keeps the characteristics unchanged, namely those of the linearized system, and the contribution due to the previous step appears solely as a driving term. In other words, the perturbative method corresponds to an integral transformation $\mathbb{T}$ obtained by integrating always along the linearized characteristic curves, instead of integrating along the characteristic curves corresponding to the previous step.

As repeatedly mentioned, the perturbation scheme has never been proved to converge but, in the case of the laser problem, the results are indistinguishable from those obtained by means of the convergent iterative process.

A.7 The dispersive laser problem. The above process can be extended, via a suitable redefinition of the mapping $\mathbb{T}$, to include dispersion effects for the linear permittivity and the nonlinear polarizability. In its simplest formulation, the dispersive laser problem consists in solving, instead of (6.193), the nonlinear integro−differential system

$$\frac{\partial H}{\partial z} = (\epsilon - 2\alpha E) \frac{\partial E}{\partial t} + \frac{\partial \mathcal{D}}{\partial t} \ , \quad \frac{\partial E}{\partial z} = \mu_o \frac{\partial H}{\partial t} \tag{6.198}$$

with the same impedance boundary conditions (6.194) on the slab walls as before. Here $\mathcal{D} = \mathcal{D}(z,t)$ is a dispersive contribution which, by force of eqs. (6.3) and (6.74), can be written in the form

$$\mathcal{D} = \frac{1}{\sqrt{2\pi}} \int_0^\infty \tilde{\psi}(\tau) E(z,\, t-\tau) d\tau + \frac{1}{2\pi} \int_0^\infty \int_0^\infty \tilde{\alpha}(\tau_1, \tau_2) E(z,\, t-\tau_1) E(z,\, t-\tau_2) d\tau_1 d\tau_2 \tag{6.199}$$

where $\tilde{\alpha}(\tau_1, \tau_2)$ and $\tilde{\psi}(\tau)$ and are the appropriate components of the memory tensor functions $\tilde{\psi}_{ijk}(\tau_1, \tau_2)$ and $\tilde{\epsilon}_{ij}(\tau) - \sqrt{2\pi}\epsilon\delta(\tau)\delta_{ij}$, respectively (see §§ 6.4 and 6.7).

Consider the sequence of linear non-dispersive problems

$$\frac{\partial H^{(n)}}{\partial z} = (\epsilon - 2\alpha E^{(n-1)})\frac{\partial E^{(n)}}{\partial t} + \frac{\partial \mathcal{D}^{(n-1)}}{\partial t}\,, \quad \frac{\partial E^{(n)}}{\partial z} = \mu_o \frac{\partial H^{(n)}}{\partial t}$$

$$\sqrt{\frac{\epsilon_o}{\mu_o}}\, E^{(n)}(0,t) - H^{(n)}(0,t) = \Psi(t) \quad , \quad \sqrt{\frac{\epsilon_o}{\mu_o}} E^{(n)}(a,t) + H^{(n)}(a,t) = 0$$

where

$$\mathcal{D}^{(n)} := \frac{1}{\sqrt{2\pi}} \int_0^\infty \tilde{\psi}(\tau) E^{(n)}(z,\, t-\tau) d\tau + \frac{1}{2\pi} \int_0^\infty \int_0^\infty \tilde{\alpha}(\tau_1, \tau_2) E^{(n)}(z,\, t-\tau_1) E^{(n)}(z,\, t-\tau_2) d\tau_1 d\tau_2$$

$(n = 1, 2, ...)$. This sequence can be envisaged as a particular case of (6.191) and (6.192), by a suitable redefinition of the source terms W_i . On the strength of previous considerations (see in particular §6.2) we may safely assume that

$$\tilde{\psi}(\tau) \in L^1(\mathbb{R}^+) \quad , \quad \tilde{\alpha}(\tau_1, \tau_2)^1(\mathbb{R}^+ \times \mathbb{R}^+) \tag{6.200}$$

where $\mathbb{R}^+ = [0, +\infty)$. Moreover, we suppose that all assumptions of Theorem 6.7.6 are satisfied.

Theorem 6.7.9 *[4]. Under the stated assumptions,if a is small enough the sequence $(E^{(n)}, H^{(n)})$ converges uniformy on $D_a \times \mathbb{R}$ to the solution (E, H) of the nonlinear dispersive problem (6.194), (6.198).*

We thus see that the nonlinear dispersive problem can be approximated by a sequence of non-dispersive linear problems. In general, it turns out that the corrections in the second-harmonic wave due to dispersion are very small (of order $O(\delta^2)$).

Remark 17. Even if dispersion in the crystal is taken into account, the boundary conditions on the slab walls remain the same, because they are dictated solely by the wave impedances of the electromagnetic field in the non-dispersive medium outside the crystal.

Exercises

Exercise 1. Let $\phi(t)$ denote a test function in the Schwartz class [2]. The inversion theorem for the Fourier transform gives the identity

$$\phi(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t}\, d\omega \int_{-\infty}^{+\infty} e^{-i\omega\tau}\ \phi(\tau)\ d\tau = \int_{-\infty}^{+\infty} \phi(\tau)\ d\tau \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega(t-\tau)}\, d\omega$$

We conclude that the Fourier transform of the Dirac distribution is $1/\sqrt{2\pi}$:

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t}\, d\omega \tag{E1}$$

Exercise 2. We must calculate the integral

$$\tilde{\psi}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \psi(\omega)\, d\omega$$

where

$$\psi(\omega) = -\frac{Ne^2}{m} \frac{1}{\omega^2 - \omega_o^2 - i\omega g} \equiv -\frac{Ne^2}{m} \frac{1}{(\omega - \omega_+)(\omega - \omega_-)} \tag{E2}$$

and

$$\omega_\pm := i\frac{g}{2} \pm \sqrt{\omega_o^2 - g^2/4}$$

Let Γ_R denote a semicircle of radius R centered at the origin in the upper-half complex $\omega-$plane for $t > 0$, in the lower-half plane for $t < 0$. Jordan's Lemma says that

$$\lim_{R\to\infty} \int_{\Gamma_R} e^{i\omega t}\ \psi(\omega) d\omega = 0$$

so that

$$\tilde{\psi}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \tilde{\psi}(\omega)\, d\omega + \lim_{R\to\infty} \int_{\Gamma_R} e^{i\omega t} \psi(\omega) d\omega$$

where $\psi(\omega)$ is given by (E2), and the integral now runs over a closed path in the complex plane ω. The function $\psi(\omega)$ is holomorphic in the lower-half plane. Hence, by applying the residue theorem as in the proof of Proposition 6.1.2 (eq. (6.22)), we find

$$\tilde{\psi}(t) \equiv 0 \qquad \text{for} \;\; t < 0$$

and

$$\tilde{\psi}(t) = \sqrt{2\pi} i (e^{i\omega_+ t}\, \mathfrak{R}_+ + e^{i\omega_- t}\, \mathfrak{R}_-) \quad \text{for } t > 0$$

where $\mathfrak{R}_\pm$ are the residues of $\psi(\omega)$ at the poles $\omega_\pm$:

$$\mathfrak{R}_+ = -\frac{Ne^2}{m} \frac{1}{\omega_+ - \omega_-} = -\frac{Ne^2}{2m} \frac{1}{\sqrt{\omega_o^2 - g^2/4}} \equiv -\mathfrak{R}_-$$

It follows that

$$\text{(E3)} \qquad \tilde{\psi}(t) = \frac{\sqrt{2\pi} N e^2}{m \sqrt{\omega_o^2 - g^2/4}} e^{-gt/2} sin\, (\sqrt{\omega_o^2 - g^2/4}\, t) \;\; \text{for } t > 0$$

which coincides with (6.38).

Exercise 3. Show that for $g > 2\omega_o$ the function $Re\psi(\omega)$ given by eq. (6.37) is decreasing for $0 < \omega^2 < \omega_o(\omega_o + g)$ and increasing for $\omega^2 > \omega_o(\omega_o + g)$.

Exercise 4. For a conductor

$$\gamma(\omega) = \frac{N'e^2}{m} \frac{1}{g + i\omega} \equiv -i \frac{N'e^2}{m} \frac{1}{\omega - ig}$$

is holomorphic for $Im\, \omega < 0$ and has a simple pole at $\omega_+ = ig$ with residue

$$\mathfrak{R}_+ = -i \frac{N'e^2}{m}$$

Jordan's Lemma implies that

$$\tilde{\gamma}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \gamma(\omega)\, d\omega + \lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{\Gamma_R} e^{i\omega t} \gamma(\omega)\, d\omega$$

and by force of the theorem of residues we have

$$\tilde{\gamma}(t) = 0 \;\text{ for } t < 0 \;; \quad \tilde{\gamma}(t) = \sqrt{2\pi} i e^{i\omega_+ t}\,\mathfrak{R}_+ = \sqrt{2\pi}\frac{N' e^2}{m} e^{-gt} \quad \text{for } t > 0$$

Exercise 5. Suppose

$$\tilde{\mathbf{E}}(\mathbf{x}, \omega) = \sqrt{2\pi}\boldsymbol{E}(\mathbf{x})\frac{\mathbb{I}_{2h}(\omega - \overline{\omega})}{2h} \quad , \quad \tilde{\mathbf{D}}(\mathbf{x}, \omega) = \tilde{\epsilon}(\omega)\sqrt{2\pi}\mathbf{E}(\mathbf{x})\frac{\mathbb{I}_{2h}(\omega - \overline{\omega}))}{2h}$$

where $\mathbb{I}_{2h}(\omega - \overline{\omega})$ is the characteristic function of the interval $\overline{\omega} - h < \omega < \overline{\omega} + h$. Then

$$\boldsymbol{E}(\boldsymbol{x}, t) = \frac{\mathbf{E}(\boldsymbol{x})}{2h}\int_{\overline{\omega}-h}^{\overline{\omega}+h} e^{i\omega t}\, d\omega = \mathbf{E}(\mathbf{x}) e^{i\overline{\omega} t}\frac{sin(ht)}{ht}$$

Since $sin(ht)\,/ht \;\cong\; 1$ for $|t| \ll h^{-1}$, the field is approximately monochromatic with frequency $\overline{\omega}$ for a finite time interval, the larger the smaller h. (This is a particular case of the sampling theorem for the Fourier transform.)

Exercise 6. Show that the limit

$$w - \lim_{h\to 0}\frac{\mathbb{I}_{2h}(\omega - \overline{\omega})}{2h} = \delta(\omega - \overline{\omega})$$

holds in the distributional (or weak) sense.

Exercise 7. Derive and discuss the Fresnel equation for a uniaxial crystal. Hint: use the fact that

$$p_3 E_3 = -\frac{\epsilon_\perp}{\epsilon_\parallel}(p_1 E_1 + p_2 E_2)$$

The Fresnel equation implies that (i) if E_3 is not zero then either $\frac{\epsilon_\perp}{\epsilon_\parallel}(p_1^2 + p_2^2) + p_3^2 = \omega^2 \epsilon_\perp \mu$ or $p^2 = \omega^2 \epsilon_\parallel \mu$, (ii) if $E_3 = 0$ then $p^2 = \omega^2 \epsilon_\perp \mu$. Geometrically, the Fresnel equation defines a surface made up of two spheres and an ellipsoid. The normal to the surface at each point coincides with $\boldsymbol{n}$, and the distance from the center is related to the refractive index [35].

Exercise 8. Perform the calculations leading to (6.63)-(6.65).

Exercise 9. Define the time variable $\tau = c_o t$ and the Riemann invariants U, V

$$U := \frac{1}{2}\big[E_2 + \sqrt{\mu_o/\epsilon_o}\, H_3\big] \quad , \qquad V := \frac{1}{2}\big[E_2 - \sqrt{\mu_o/\epsilon_o}\, H_3\big]$$

(cfr. Exercise 14 of Chapter 4). The linear Maxwell equations (6.94) are easily seen to be equivalent to the relations

$$\text{(E4)} \qquad \frac{\partial U}{\partial \tau} + \frac{\partial U}{\partial x} = 0 \ , \quad \frac{\partial V}{\partial \tau} - \frac{\partial V}{\partial x} = 0$$

which imply that $U = U(t-\tau)$ is an arbitrary function of $(t - x/c_o)$, invariant along the characteristic lines $t - x/c_o = \mathit{constant}$, and $V = V(t+\tau)$ is an arbitrary function of $(t + x/c_o)$, invariant along the characteristic lines $t - x/c_o$ =constant. Hence

$$E_2 = U(t - x/c_o) + V(t + x/c_o)\,, \quad H_3 = \sqrt{\epsilon_o/\mu_o}\left[U(t - x/c_o) - V(t + x/c_o)\right]$$

For $x < 0$ the invariants U, V must coincide with the appropriate components of the incident and reflected waves, respectively, whereas for $x > a$ the invariant U coincides with the appropriate component of the transmitted wave and $V \equiv 0$, since there is no wave "reflected from infinity".

Similar conclusions follow from eqs. (6.95) by defining the Riemann invariants

$$U := \frac{1}{2}\left[E_3 - \sqrt{\mu_o/\epsilon_o}\, H_2\right] \quad , \qquad V := \frac{1}{2}\left[E_3 + \sqrt{\mu_o/\epsilon_o}\, H_2\right]$$

so that U and V satisfy (E4). Hence

$$E_3 = U(t - x/c_o) + V(t + x/c_o) \quad , \quad H_3 = \sqrt{\epsilon_o/\mu_o}\ \left[-U(t - x/c_o) + V(t + x/c_o)\right]$$

where the invariant U is related to the incident wave (for $x < 0$) or to the transmitted wave (for $x > a$), whereas V is related to the reflected wave and vanishes for $x > a$.

Exercise 10. Hint: Look for solutions in the (complex) form

$$E_3(x,t) = Re\left[A_\parallel exp(i\omega t\ - ip_\parallel x) + B_\parallel exp(i\omega t + ip_\parallel x)\right] cos\,\theta$$
$$H_2(x,t) = \sqrt{\epsilon_\parallel/\mu_o}\, Re\left[C_\parallel exp(i\omega t - ip_\parallel x) + D_\parallel exp(i\omega t + ip_\parallel x)\right] cos\,\theta$$

Substituting in the differential equations (6.104) we find $C_\parallel = -A_\parallel$, $D_\parallel = B_\parallel$. The boundary conditions (6.105) yield eqs. (6.108). The uniqueness theorem then guarantees that (6.109) is the required (bounded) solution.

Exercise 11. Hint: Let t_o denote a delay time defined by the relation

$$e^{i\omega t_o} = \frac{1 - r_\perp^2 exp(2ip_\perp a)}{\sqrt{d_\perp}}$$

The first eq. (6.114) takes then the form

$$E_o(x,t) = \frac{2E_l sin\,\theta}{(n_{or}+1)\sqrt{d_\perp}}\, Re\big[e^{i\omega(t+t_o)-ip_\perp x} + r_\perp e^{i\omega(t+t_o)+ip_\perp(x-2a)}\big]$$
$$= \frac{2E_l sin\,\theta}{(n_{or}+1)\sqrt{d_\perp}}\Big\{cos\big[\omega(t+t_o)-p_\perp x\big] + r_\perp cos\big[\omega(t+t_o)+p_\perp(x-2a)\big]\Big\}$$

whence

$$\frac{\partial E_o}{\partial t} = \frac{-2E_l\,\omega sin\,\theta}{(n_{or}+1)\sqrt{d_\perp}}\Big\{sin\big[\omega(t+t_o)-p_\perp x\big] + r_\perp sin\big[\omega(t+t_o)+p_\perp(x-2a)\big]\Big\}$$

Thus in eq. (6.122) we find that the forcing term $E_o\frac{\partial E_o}{\partial t}$ is periodic with period π/ω :

$$E_o\frac{\partial E_o}{\partial t} = \frac{-4E_l^2\,\omega sin^2\theta}{(n_{or}+1)^2 d_\perp}\Big\{\frac{1}{2}sin\big[2\omega(t+t_o)-2p_\perp x\big] + \frac{1}{2}r_\perp^2 sin\big[2\omega(t+t_o)+2p_\perp(x-2a)\big]$$
$$+\ r_\perp sin\big[2\omega(t+t_o)-2p_\perp a)\big]\Big\}$$

and this suggests to look for solutions of (6.120) and (6.122) of the (complex) form (6,123) and (6.124). Indeed, the complex constants $k_o, ..., k_8$ turn out to be uniquely determined. In particular for the second-harmonic waves one finds

$$k_7 = \frac{n_{or}}{(n_{or}+1)^3 d_\perp}\Big\{[1 + r_\perp e^{-4ip_\perp a}] + \frac{4ip_\perp a r_\perp(p_\perp+1) + 4r_\perp[e^{2ip_\perp a} - r_\perp e^{-2ip_\perp a}]}{exp(4ip_\perp a) - r_\perp^2}$$
$$- \frac{8r_\perp + 2n_{or}(1-r_\perp^2)}{(n_{or}+1)(exp(4ip_\perp a) - r_\perp^2)}\Big\}$$

and

$$k_8 = e^{2i(p_o-p_\perp)a}k_1 + e^{2i(p_o+p_\perp)a}k_2 + \frac{2e^{2i(p_o-p_\perp)a}[ip_\perp a\,(1-r_\perp^2) + 2r_\perp]}{(n_{or}+1)^2 d_\perp}$$

where

$$k_1 = \frac{1}{2}(1-n_{or}^{-1})k_7 - \frac{2r_\perp e^{-2ip_\perp a} + \frac{1}{2}r_\perp^2 e^{-4ip_\perp a} - \frac{1}{2}}{(n_{or}+1)^2 d_\perp}$$

$$k_2 = \frac{1}{2}(1 + n_{or}^{-1})k_7 - \frac{2r_\perp e^{-2ip_\perp a} - \frac{1}{2} r_\perp^2 e^{-4ip_\perp a} + \frac{1}{2}}{(n_{or}+1)^2 d_\perp}$$

The remaining constants are given by [6]

$$k_3 = \frac{2ip_\perp}{(n_{or}+1)^2 d_\perp} \ , \ k_4 = -\frac{2ip_\perp r_\perp^2 e^{-4ip_\perp a}}{(n_{or}+1)^2 d_\perp} \ , \ k_o = \frac{4r_\perp e^{-2ip_\perp a}}{(n_{or}+1)^2 d_\perp}$$

and

$$k_5 = k_1 - \frac{1}{(n_{or}+1)^2 d_\perp} \quad , \quad k_6 = -k_2 + \frac{r_\perp^2 e^{-4ip_\perp a}}{(n_{or}+1)^2 d_\perp}$$

Exercise 12. Write the explicit expressions of the reflected and transmitted waves for propagation along the $y-$axis (§6.7.5).

Exercise 13. Show that

$$E_o(x,t) = Re\big[A_\perp exp(i\omega t - ip_\perp x) + B_\perp exp(i\omega t + ip_\perp x)\big]$$
$$H_o(x,t) = \sqrt{\epsilon_\perp/\mu_o}\, Re\big[A_\perp exp(i\omega t - ip_\perp x) - B_\perp exp(i\omega t + ip_\perp x)\big]$$

where $A_\perp$, $p_\perp$, $B_\perp$ are defined in §6.7.4. Hence

$$E_o = \frac{2E_l}{(n_{or}+1)d_\perp}\big[cos(i\omega t - ip_\perp z) - r_\perp^2 cos(i\omega t - ip_\perp z + 2ip_\perp a)$$
$$+ r_\perp cos(i\omega t + ip_\perp z - 2ip_\perp a)\big]$$

$$\frac{\partial E_o}{\partial t}(z,t) = \frac{-2E_l\omega}{(n_{or}+1)d_\perp}\big[sin\,(i\omega t - ip_\perp z) - r_\perp^2 sin(i\omega t - ip_\perp z + 2ip_\perp a)$$
$$+ r_\perp sin(i\omega t + ip_\perp z - 2ip_\perp a)\big]$$

and $E_o \partial E_o/\partial t$ is a superposition of sinusoidal functions of (circular) frequency 2ω.

Exercise 14. Show that

$$E'(z,t) = E_l Re\Big[e^{2i\omega t}\big\{k_o + k_1 e^{-2ip_\perp z} + k_2 e^{2ip_\perp z} + k_3 z e^{-2ip_\perp z} + k_4 x e^{2ip_\perp z}\big\}\Big]$$
$$H'(z,t) = -E_l\sqrt{\frac{\epsilon_\perp}{\mu_o}}\, Re\Big[e^{2i\omega t}\big\{k_5 e^{-2ip_\perp z} + k_6 e^{2ip_\perp z} + k_3 z e^{-2ip_\perp z} - k_4 z e^{2ip_\perp z}\big\}\Big]$$

and the second-harmonic reflected and transmitted waves are given by

$$\tilde{\mathcal{R}}_1(t) = \tilde{\mathcal{R}}(t) sin2\theta,\ \tilde{\mathcal{R}}_2(t) = \tilde{\mathcal{R}}(t)\, cos2\theta\,, \tilde{\mathcal{T}}_1(t) = \tilde{\mathcal{T}}(t) sin2\theta\,, \tilde{\mathcal{T}}_2(t) = \tilde{\mathcal{T}}(t)\, cos2\theta$$

where

$$\tilde{\mathcal{R}}(t) := \delta E_l \, Re\Big[k_7 \, e^{2i\omega t}\Big] \quad , \qquad \tilde{\mathcal{T}}(t) := \delta E_l Re\Big[k_8 e^{2i\omega t}\Big]$$

and $k_o, ..., k_8$ are the complex constants defined previously (see Exercise 11).

Exercise 15. Prove that eq. (4.122) for the bicharacteristic rays reduces to the solutions of eq. (6.165) when the system matrices are given by $\mathbb{A}_k := \mathbb{S}^{-1}\mathbb{D}_k\mathbb{S}$.

Exercise 16. Show that the hyperbolic system (6.146) can be written by purely algebraic manipulations in the characteristic form

$$\sum_{j=1}^{4} S_{ij}(z,t,\boldsymbol{u})\Big[\frac{\partial u_j}{\partial z} + \rho_i(z,t,\boldsymbol{u})\frac{\partial u_j}{\partial t}\Big] = W_i(z,t,\boldsymbol{u}) \quad (i = 1, ..., 4)$$

where $\boldsymbol{u}= (\tilde{E}_1, \tilde{H}_2, \tilde{E}_2, \tilde{H}_1)$, ρ_i are the four eigenvalues given by

$$\rho_i = (-1)^{i+1}\frac{\sqrt{1 + 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}}{c_\perp} \quad , \quad \rho_{i+2} = (-1)^i \frac{\sqrt{1 - 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}}{c_\perp}$$

$(i = 1, 2)$. Here $|\boldsymbol{E}(z,t)| := \sqrt{E_1^2 + E_2^2}$ and the functions $S_{ij}(z,t,\boldsymbol{u})$, $\rho_i(z,t,\boldsymbol{u})$ are regular in a suitable neighborhood Ω of $\boldsymbol{u}= \boldsymbol{0}$ for all $0 \le z \le a$, $t \in \mathbb{R}$.

Prove that the matrix (S_{ij}) is not diagonally dominant.

Hint: For E_j^o $(j = 1, 2)$ given as the functions of (z,t) obtained by solving (6.143) and (6.145), we can rewrite the hyperbolic system (6.146) in the vector form

$$\boldsymbol{u}_z + \mathbb{A}(z,t,\boldsymbol{u})\boldsymbol{u}_t = \boldsymbol{w}(z,t,\boldsymbol{u})$$

where $\boldsymbol{u}= (\tilde{E}_1, \tilde{H}_2, \tilde{E}_2, \tilde{H}_1)$ and $\boldsymbol{w}= (0, \mathcal{F}_1, 0, \mathcal{F}_2)$. The matrix $\mathbb{A}$ has four independent left eigenvectors $\boldsymbol{h}_i$ corresponding to ρ_i, $i = 1, ..., 4$, which can be determined globally in Ω, where Ω is typically the hypercube $|\mathrm{u}| < M$, $M > 0$. Let

$$\mathbb{S} = \begin{pmatrix} h_1^1 & h_1^2 & h_1^3 & h_1^4 \\ h_2^1 & h_2^2 & h_2^3 & h_2^4 \\ h_3^1 & h_3^2 & h_3^3 & h_3^4 \\ h_4^1 & h_4^2 & h_4^3 & h_4^4 \end{pmatrix}$$

Then

$$\mathbb{S}\mathbb{A} = \mathbb{D}\mathbb{S} \quad , \quad \mathbb{D} = \begin{pmatrix} \rho_1 & 0 & 0 & 0 \\ 0 & \rho_2 & 0 & 0 \\ 0 & 0 & \rho_3 & 0 \\ 0 & 0 & 0 & \rho_4 \end{pmatrix}$$

and the system takes the form

$$\mathbb{S}\boldsymbol{u}_z + \mathbb{D}\mathbb{S}\boldsymbol{u}_t = \mathbb{S}\boldsymbol{w}(z,t,\boldsymbol{u})$$

The eigenvectors have the form

$$\boldsymbol{h}_i = (c_\perp \rho_i (E_2 \pm |\boldsymbol{E}|),\ (E_2 \pm |\boldsymbol{E}|),\ c_\perp \rho_i E_1,\ -E_1)$$

($+$ for $i = 1,2$, $-$ for $i = 3,4$) and thus are regular functions in Ω. However, the matrix $\mathbb{S}$ is not diagonally dominant.

Exercise 17. In Exercise 16, take the new variables $\boldsymbol{U} = \mathbb{L}(\boldsymbol{u})$ defined by the linearized Riemann invariants

$$U_1(z,\tau) = \frac{1}{2}\left[\sqrt{\frac{\epsilon_\perp}{\mu_o}}\, E_1(z,t) + H_2(z,t)\right] \quad , \quad U_2(z,\tau) = \frac{1}{2}\left[\sqrt{\frac{\epsilon_\perp}{\mu_o}}\, E_1(z,t) - H_2(z,t)\right]$$

$$U_3(z,\tau) = \frac{1}{2}\left[\sqrt{\frac{\epsilon_\perp}{\mu_o}} E_2(z,t) + H_1(z,t)\right] \quad , \quad U_4(z,\tau) = \frac{1}{2}\left[\sqrt{\frac{\epsilon_\perp}{\mu_o}}\, E_2(z,t) - H_1(z,t)\right]$$

where $\tau = c_\perp t$. Show that:

(i) The matrix $\mathbb{S}$ of the characteristic form of the system (6.146) in these new variables is

$$\text{(E5)} \quad \mathbb{S}(z,\tau,\boldsymbol{U}) = \begin{pmatrix} 1 & \mathrm{O}(\delta) & \mathrm{O}(\delta) & cotan\frac{\theta}{2} + \mathrm{O}(\delta) \\ \mathrm{O}(\delta) & 1 & \mathrm{O}(\delta) & \mathrm{O}(\delta) \\ \mathrm{O}(\delta) & \mathrm{O}(\delta) & 1 & \mathrm{O}(\delta) \\ tan\frac{\theta}{2} + \mathrm{O}(\delta) & \mathrm{O}(\delta) & \mathrm{O}(\delta) & 1 \end{pmatrix}$$

(ii) This matrix is not diagonally dominant as $\delta \to 0$ (for any value of the polarization θ)

(iii) For $\delta = 0$ the functions $U_1, ..., U_4$ should satisfy the diagonal system

$$\frac{\partial U_i}{\partial z} + \rho_i \frac{\partial U_i}{\partial \tau} = 0 \ , \ \ i = 1, ..., 4$$

with $\rho_1 = \rho_4 = 1$, $\rho_2 = \rho_3 = -1$. However, not all of these equations are recovered correctly from (E5) in the limit $\delta \to 0$, since in this limit the first and the fourth equation reduce to the single equation

$$sin\frac{\theta}{2}\left(\frac{\partial U_1}{\partial z} + \frac{\partial U_1}{\partial \tau}\right) + cos\frac{\theta}{2}\left(\frac{\partial U_4}{\partial z} + \frac{\partial U_4}{\partial \tau}\right) = 0$$

which is unable to determine both U_1 and U_4.

Exercise 18. Show that the characteristic curves of (6.144) are given by

$$z = c_j^{\pm} t + constant \qquad (j = 1, 2)$$

with four characteristic speeds $c_1^+, c_2^+, c_1^-, c_2^-$ depending on $|\boldsymbol{E}| := \sqrt{E_1^2 + E_2^2}$:

$$\text{(E6)} \qquad c_1^{\pm} = \pm \frac{c_\perp}{\sqrt{1 + 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}} \quad , \quad c_2^{\pm} = \pm \frac{c_\perp}{\sqrt{1 - 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}}$$

so that the four characteristic speeds coalesce pairwise $c_1^+ = c_2^+$, $c_1^- = c_2^-$ at points (z,t) where $\alpha |\boldsymbol{E}(z,t)| = 0$ (in particular, everywhere for $\alpha = 0$).

Hint: if $\boldsymbol{u} := (E_1, H_2, E_2, H_1)$, the system (6.144) can be written in vector form as

$$\boldsymbol{u}_z + \mathbb{A}\boldsymbol{u}_t = \boldsymbol{0} \quad , \quad \mathbb{A} = \begin{pmatrix} 0 & \mu_o & 0 & 0 \\ \epsilon_\perp + 2\alpha E_2 & 0 & 2\alpha E_1 & 0 \\ 0 & 0 & 0 & -\mu_o \\ -2\alpha E_1 & 0 & -\epsilon_\perp + 2\alpha E_2 & 0 \end{pmatrix}$$

The four roots

$$\rho_i = (-1)^{i+1} \frac{\sqrt{1 + 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}}{c_\perp} \quad , \quad \rho_{i+2} = (-1)^{i} \frac{\sqrt{1 - 2c_\perp^2 \mu_o \alpha |\boldsymbol{E}(z,t)|}}{c_\perp}$$

$(i = 1, 2)$ of the characteristic equation

$$det\,(\rho \mathbb{I} - \mathbb{A}) \equiv \rho^4 - 2\epsilon_\perp \mu_o \rho^2 + (\epsilon_\perp \mu_o)^2 - 4\mu_o^2 \alpha^2 |\boldsymbol{E}|^2 = 0$$

yield the inverse of the characteristic speeds (E6), depending on (E_1, E_2). For $E_1 = E_2 = 0$ the four distinct speeds degenerate into the two double speeds of the linearized system ($\alpha = 0$).

Exercise 19. Show that eqs. (6.184) and (6.185) imply that the energy relation holds

$$\frac{1}{2}\frac{d}{dt}\int_0^a (u^2 + v^2)dx = h\big[u^2(a,t) + u^2(0,t)\big]$$

Thus if $h > 0$ the $L^2(0,a)-$norms of u and v may diverge as $t \to +\infty$, if $h < 0$ this is may happen as $t \to -\infty$, and if $h = 0$ these norms are time-independent. Cfr. Exercise 15, Chapter 4.

References

[1]A.K. Assis, W.A. Rodrigues, A.J. Mania, The electric field outside a stationary resistive wire carrying a constant current, Found. Physics 29, 729-753, 1999

[2]P. Bassanini, A.R. Elcrat, Theory and Applications of Partial Differential Equations, Springer/Plenum Press, 1997

[3]P. Bassanini, L. Cesari, La duplicazione di frequenza nella radiazione laser, Rend. Accad,Lincei VIII, 69, 3-4, 166-173, 1980

[4]P. Bassanini, The problem of Graffi-Cesari. In: Proc. Symp. on Nonlinear Phenomena in the Math.Sciences,Arlington, TX, 1980; Academic Press, 87-101, 1982

[5]P. Bassanini, A nonlinear hyperbolic problem arising from a question of nonlinear optics, Zeitschr. Angew. Math. Phys. 27, 409-422, 1976

[6]P. Bassanini, M.C. Salvatori, Problemi ai limiti per sistemi iperbolici quasilineari e generazione di armoniche ottiche, Riv. Mat. Univ. Parma 5, 55-76, 1979

[7]P. Bassanini, M.C. Salvatori, A theorem of existence and uniqueness in nonlinear optics, Boll. UMI 16-B, 597-611, 1979

[8]G.K. Batchelor, An Introduction to Fluid Dynamics, Cambridge University Press, 1991

[9]A. Berti, Reflection-transmission problem in stratified media, Ph.D. Thesis, Univ.of Genova, Italy, 2007

[10]A. Berti, Existence and uniqueness for the reflection and transmission problem in stratified electromagnetic media, J. Math. Phys. 47, 543-564, 2006

[11]A.A. Blank, K.O. Friedrichs, H. Grad, Theory of Maxwell's equations without displacement current, Report NYO-6486, Inst. Math Sciences, NYU, 1957

[12]N. Bloembergen, Nonlinear Optics, Benjamin 1977

[13]F. Brezzi, G. Gilardi, Fundamentals of PDE's for numerical analysis, Report n. 446, IAN-CNR, Pavia, Italy, 1984

[14]G. Caviglia, A. Morro, Existence and uniqueness in the reflection-transmission problem, Q. Jl.Mech. appl. Math. 52 (4), 543-564, 1999

[15]L. Cesari, A boundary value problem for quasilinear hyperbolic systems, Ann. Scuola Norm. Sup. Pisa (4) **1**, 311-358, 1974

[16]L. Cesari, Nonlinear oscillations under hyperbolic systems. In: Dynamical Systems, An International Symposium, Academic Press, 251-261, 1976

[17]B.D. Coleman, E.H. Dill,Thermodynamic restrictions on the constitutive equations of electromagnetic theory, Zeitschr. angew. Math. Phys. 26, 61-76, 1975

[18]R. Courant, D. Hilbert, Methods of Mathematical Physics II, Interscience 1964

[19]M.P. do Carmo, Differential Geometry of Curves and Surfaces, Prentice-Hall, 1976

[20]M. Fabrizio, A. Morro, Electromagnetism of continuous media, Oxford University Press, Oxford 2003

[21]M. Fabrizio, Théorèmes d' unicité sur un nouveau problème aux limites rélatif aux équations de Maxwell non linéaires, J. Mec. 15, 681-696, 1976

[22]N.H. Frank, Introduction to Electricity and Optics, McGraw-Hill, 1950

[23]P.A. Franken, A.E. Hill, C.W. Peters, G. Weinreich, Generation of Optical Harmonics, Phys. Rev. Letters 7, 118-119, 1961

[24]P.A. Franken, J. F. Ward, Optical Harmonics and Nonlinear Phenomena, Rev.Modern Phys. 35,23-39, 1963

[25]K.O. Friedrichs, Mathematical methods of electromagnetic theory, Courant Institute of Mathematical Sciences, New York, 1974

[26]D. Graffi, Nonlinear partial differential equations in physical problems, Res. Notes in Maths., Pitman 1980

[27]D. Graffi, Teoria matematica dell' Elettromagnetismo, Patron (Bologna, Italy), 1972

[28]D. Graffi, Onde elettromagnetiche, CNR, Roma, 1965

[29]D.J. Griffiths, Introduction to Electrodynamics, Benjamin-Cummings Pub. Co., 1998

[30]O. Heaviside, Electromagnetic Theory, AMS Chelsea Publishing, 1971

[31]F. John, Partial Differential Equations, Springer 1982

[32]D.S. Jones, The Theory of Electromagnetism, Pergamon Press, 1964

[33]C. Kittel, Introduction to Solid State Physics, Wiley & Sons, 1956

[34]R. Kress, Linear Integral Equations, Springer, 1994

[35]L. D. Landau, E. M. Lifshitz, L. P. Pitaevskij, Electrodynamics of continuous media, MIR,Moscow 1986

[36]J. Marsden, A. Weinstein, Calculus, Springer-Verlag, 1985

[37]E. Martensen, Potentialtheorie, Teubner, 1965

[38]N. I. Mushkelishvili, Singular Integral Equations, Dover, 1992

[39]A. W. Naylor, G.R. Sell, Linear Operators Theory in Engineering and Science, Springer 1982

[40]M. Recchioni ; F. Zirilli "A new formulation for time-dependent e.m. scattering from a bounded obstacle", J. Engrg. Math. 47, 1, 17-47, 2003

[41]S. Singh, Non-linear Optical Materials. In: C.R.C.Handbook of Lasers, CRC Press, 1971, p. 489-522

[42]J.C. Slater, N.H. Frank, Electromagnetism, McGraw-Hill, 1947

[43]A.Sommerfeld, Electrodynamics, Academic Press, 1952

[44]C. Truesdell ; R.A. Toupin, "The classical field theories", Handbuch der Physik III/1, Springer, 1960

[45]H.F. Weinberger, Partial differential equations, J. Wiley & Sons, 1965

Index