

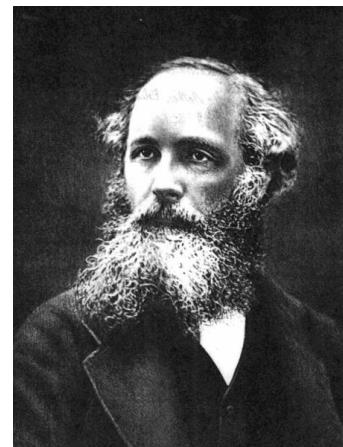
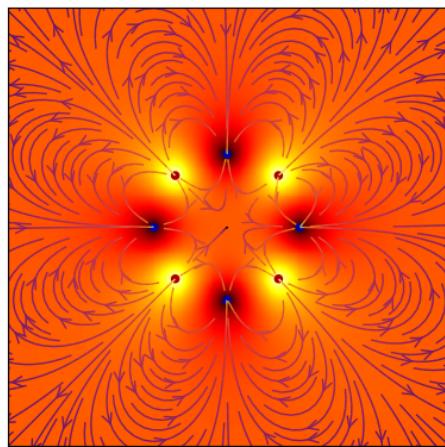
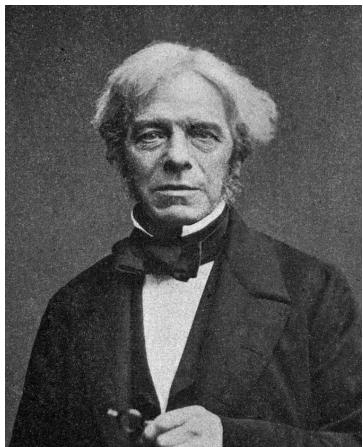


Arnaud Couairon, Fabian Cadiz

PHY204



Classical Electrodynamics



Edition 2023

Foreword

The 2023 edition of this textbook is available to the Bachelor students at Ecole Polytechnique with the goal of providing them with written support for the course of electrodynamics (PHY204).

The organization of this textbook (one chapter per lecture) reflects the organization of the course, without the content being strictly identical. A course made in person indeed requires adaptations to match the constraints of the calendar. Changes, even if minor, are usually made after the textbook is printed. The material of the textbook also contains reminders of the course of electromagnetism and light (PHY104) which are presented in more details in the textbook than they will be during the PHY204 lectures.

In this respect, the readers will note some overlap with the textbook of electromagnetism and light, partly due to the goal of providing the necessary reminders, and partly due to the fact that the very first version of this textbook was originally written by Fabian Cadiz in Spanish, translated in English, supplemented by new chapters and reorganized for the Bachelor students of both courses of electromagnetism and light and electrodynamics.

This textbook also provides an opportunity to give more details on specific applications of the concepts, for instance to modern technology (touchscreen devices, microwave oven, laser amplifiers, etc), than can be given during the lectures. The laws of electrodynamics are most naturally expressed using vectorial operators, vector calculus theorems and identities, and other mathematical tools (which also appear in other sciences, for instance fluid mechanics). A toolbox is provided in the appendices that the students can use to come back, at their own pace, on calculus details which might be partly skipped during the lectures but are thoroughly covered in the textbook.

A few examples and simple exercises illustrate the different concepts throughout the textbook. However, the students who would like to train with more exercises, including fully solved challenging exercises, may try the exercise textbook in pdf format on the PHY204 moodle page.

Several books on electrodynamics have inspired us for writing this textbook, which is in no way the only source of information for PHY204. The students who wish to look for complementary material on the topics covered in the course and dive in a much broader field can refer to:

- D.J. Griffith (2013), *Introduction to Electrodynamics*, (4th ed.) Pearson
- A. Zangwill (2013), *Modern Electrodynamics*, Cambridge
- Feynman, Leighton, Sands, *The Feynman lectures on physics*, vol II.

<http://www.feynmanlectures.caltech.edu/>

The present textbook evolves each year. The students who would like to give feedback on their expectations and use of the textbook are very welcome to talk with, or write to the authors. The future versions will take their remarks into account.

A. Couairon and F. Cadiz
Palaiseau, February 2023

Contents

1 Maxwell's equations - Electrostatics	13
1.1 Postulates of electromagnetism	14
1.1.1 Field-field Maxwell's equations	14
1.1.2 Field-source Maxwell's equations	16
1.1.3 The Lorentz force	18
1.2 Gauss's theorem in vacuum	19
1.3 Charge conservation	20
1.3.1 Current through a surface	21
1.3.2 Charge conservation law - Integral form	21
1.3.3 Charge conservation law - Differential form	22
1.4 The electrostatic field	23
1.4.1 The Laplace and Poisson equations	23
1.4.2 Boundary conditions for Poisson's equation	24
1.4.3 Solutions to the Poisson equation	24
1.4.4 The Dirac distribution, the mathematical problem of dealing with point charges	27
1.4.5 Charge and current densities for discrete charge distributions	29
1.5 Symmetry arguments in electrostatics	30
1.5.1 Invariance by spatial translation along an axis	31
1.5.2 Invariance by spatial rotation around an axis	31
1.5.3 Mirror symmetries	32
1.6 Exercises	34
1.7 Summary and essential formulas	37
2 Electrostatics of charge distributions	39
2.1 General expressions for the field and potential generated by a charge distribution	40
2.2 Multipole expansion of the potential for a discrete charge distribution	42
2.3 Multipole expansion of the potential for a continuous charge distribution	47
2.4 Expansion in Legendre polynomials	50
2.4.1 Multipole expansion with spherical coordinates	51
2.5 Energy for a discrete charge distribution	53
2.5.1 Energy of a single charge in a preexisting field	53
2.5.2 Constitution energy for a discrete charge distribution	54

2.5.3	Constitution energy for a continuous charge distribution	55
2.5.4	Electrostatic energy density	55
2.6	Effect of an external electric field on a dipole distribution	56
2.7	Exercises	59
2.8	Summary and essential formulas	63
3	Metals, conductors and capacitors	65
3.1	Model for metals	66
3.1.1	Fixed and mobile charges	66
3.1.2	Average density of carriers	67
3.1.3	Spatially averaged electric fields	68
3.2	The Drude model for electric conductivity	68
3.3	Conductors in static equilibrium	70
3.3.1	The electric field in a conductor at equilibrium is zero	70
3.3.2	The charge of a conductor in equilibrium is distributed on its surface	70
3.3.3	The electric field is normal to the surface of a conductor in equilibrium	71
3.3.4	Coulomb's theorem: electric field at the surface of a conductor in equilibrium	72
3.3.5	Electrostatic energy of a conductor	73
3.4	Interaction between conductors	73
3.4.1	Electrostatic induction	73
3.4.2	Capacitance	75
3.4.3	Generalization: induction coefficient between conductors	76
3.5	The capacitor	77
3.5.1	Capacitance of a capacitor	78
3.5.2	Electrostatic energy of a capacitor	78
3.5.3	Determination of a capacitance	79
3.6	Force on a charged conductor: the electrostatic pressure	83
3.6.1	Force on a capacitor	84
3.7	Thermodynamic potentials for the forces acting on a conductor	88
3.7.1	Free energy of a capacitor	88
3.7.2	Forces on a conductor by means of the free energy	90
3.7.3	Another thermodynamic potential: the free enthalpy	91
3.8	Summary and essential formulas	93
4	The field in dielectric media	95
4.1	Polarization vector	96
4.1.1	Potential generated by a dielectric medium	97
4.1.2	Bound charge densities	97
4.2	Gauss's law in a dielectric	99
4.3	Electric susceptibility and dielectric constant	100
4.3.1	Linear homogeneous and isotropic media	101
4.3.2	Electric susceptibility and dielectric constant	102
4.3.3	Physical interpretation of the dielectric constant	103

4.4	Boundary conditions at the interface between dielectric media	106
4.4.1	The normal component of the displacement field is discontinuous at a charged interface	106
4.4.2	The tangential component of the electric field is continuous at an interface	107
4.5	Capacitors with dielectrics	108
4.5.1	Capacity of a capacitor with dielectric medium	108
4.5.2	Capacitance coefficients in a dielectric medium	109
4.6	The electrostatic free energy for a polarized medium	111
4.6.1	Electrostatic free energy density for a dielectric inserted in a plane capacitor	111
4.6.2	Free energy density for a dielectric medium - general case	112
4.6.3	Free energy density for a linear homogeneous and isotropic medium . . .	113
4.6.4	Free enthalpy	114
4.7	Forces applied to a dielectric medium	115
4.7.1	Force and torque density	115
4.7.2	Electrostatic pressure for a linear homogeneous isotropic medium . . .	116
4.8	Gauss's law: from vacuum to a dielectric	117
4.9	Summary and essential formulas	119
5	Dielectrics: Microscopic study and effect of time-dependent fields	121
5.1	Microscopic study of dielectrics	121
5.1.1	The depolarizing field	122
5.1.2	The local field	123
5.2	Polarization mechanisms	126
5.2.1	Electronic polarization	126
5.2.2	Ionic polarization	127
5.2.3	Polarization by dipolar orientation	129
5.2.4	Clausius-Mossotti's relation	131
5.3	Dielectrics in time-dependent fields	132
5.3.1	The Lorentz model for the classical atom (1878)	132
5.3.2	Orientation polarization in the sinusoidal regime	137
5.3.3	The complex permittivity - dielectric heating	139
5.4	Summary and essential formulas	142
6	Metal conductivity, conductors and electric currents, Hall effect	143
6.1	Reminders about conductors in static equilibrium	144
6.2	Microscopic description of conductivity	145
6.2.1	Fixed and mobile charges	145
6.2.2	Average density of carriers	145
6.2.3	Average electric field	146
6.2.4	Spatially averaged electric fields	146
6.2.5	Collisions in a real metal	147
6.2.6	Carrier lifetime	148
6.2.7	Motion of carriers	149
6.3	Electric currents	150

6.4	Ohm's law - Microscopic form	152
6.4.1	Ohm's law in the case of a time-dependent electric field	153
6.4.2	About the validity of the Drude model	154
6.5	Ohm's law - Macroscopic form	155
6.5.1	Relation between resistance and capacity	157
6.5.2	Charge neutrality in a conductor	158
6.6	Electric energy and power	159
6.6.1	Power supplied by a battery	160
6.6.2	Joule effect (1860) - Power dissipated in a resistance	160
6.7	The Hall effect	160
6.7.1	Applications of the Hall effect	164
6.8	Summary and essential formulas	166
7	Magnetostatics	169
7.1	The magnetic field	170
7.1.1	The magnetic force	170
7.1.2	Where does the magnetic field of a magnet come from?	171
7.1.3	Lorentz Force	171
7.2	Field from a current - The Biot-Savart law (1820)	171
7.2.1	Magnetic field from a continuous current distribution	172
7.2.2	Magnetic field generated by a line current	173
7.2.3	Magnetic field lines	173
7.3	Magnetic fields for common current distributions	174
7.4	Ampère's law	175
7.4.1	Ampère's law - Differential form	176
7.4.2	Ampère's law - Integral form	177
7.5	Symmetry arguments in magnetostatics	178
7.6	The vector potential	180
7.6.1	Vector potential of a point charge	181
7.6.2	Vector potential of an arbitrary current distribution	181
7.6.3	The vector potential is not unique	182
7.6.4	Poisson's equation for the vector potential	182
7.7	The magnetic dipole	183
7.7.1	Magnetic field of a magnetic dipole	183
7.7.2	Magnetic moment	185
7.7.3	Multipolar expansion for the magnetic field	189
7.8	Forces exerted on a magnetic dipole in an external magnetic field	191
7.8.1	The Laplace force	191
7.8.2	Force on a magnetic dipole	193
7.8.3	Torque on a magnetic dipole	195
7.9	The two fundamental laws of magnetostatics	196
7.9.1	Gauss's law for magnetism	196
7.9.2	The two fundamental laws of magnetostatics	197

7.10	Completeness of magnetostatics	197
7.11	Summary and essential formulas	198
8	Faraday's law of induction	201
8.1	Induced currents: experimental observations	202
8.1.1	Deformation of a circuit in a static magnetic field	202
8.1.2	Mobile circuit in a static magnetic field	203
8.1.3	Circuit at rest in a time-varying magnetic field	203
8.1.4	Lenz's law (1834)	204
8.2	Electromotive force - Faraday's law	204
8.2.1	Magnetic flux through a moving surface	205
8.2.2	Mobile circuit in a static magnetic field	206
8.2.3	Static circuit in a time-varying magnetic field	208
8.3	Faraday's law, differential form	209
8.3.1	Generalized Ohm's law	211
8.4	Inductance	213
8.4.1	Self-inductance	213
8.4.2	Magnetic coupling between current loops	216
8.4.3	Energy stored in an inductance	219
8.5	Magnetostatic energy	220
8.5.1	Energy function for a filamentary circuit in an external magnetic field	220
8.5.2	Forces and torque on a loop in a magnetic field	222
8.5.3	Magnetic free energy	223
8.5.4	Another expression for the volume density of magnetic free energy	227
8.5.5	Free energy vs free enthalpy	228
8.6	Reminders on energy and potentials in thermodynamics	231
8.7	Summary and essential formulas	235
9	Magnetism in matter	239
9.1	The different types of magnetism	239
9.2	Microscopic origin of magnetism	239
9.3	Dynamics of a magnetic dipole moment in a magnetic field	243
9.3.1	The Larmor precession	244
9.4	Magnetization of matter	247
9.4.1	The microscopic origin of diamagnetism	248
9.4.2	Microscopic origin of Paramagnetism and Ferromagnetism	248
9.4.3	From electronics to spintronics	249
9.5	Field by a magnetized material	250
9.5.1	The magnetization current densities	251
9.6	Magnetostatics in a magnetized medium	253
9.7	Constitutive relations of a magnetic material	254
9.8	Hysteresis	258
9.9	Boundary conditions for the magnetic field	261
9.9.1	The normal component of the magnetic field is continuous at an interface	262

9.9.2	The tangential component of the magnetic excitation is discontinuous at a current carrying surface	262
9.10	The magnetic free energy	264
9.10.1	Differential expression for the volume density of free energy	264
9.10.2	Properties for the magnetic free energy	266
9.10.3	Free energy vs free enthalpy	266
9.10.4	Free energy for linear, homogeneous, isotropic media	266
9.11	Summary and essential formulas	268
10	Maxwell's equations and wave propagation in matter	271
10.1	Ampère's theorem and displacement current	272
10.1.1	Displacement current	273
10.2	D'Alembert's wave equation	276
10.2.1	Plane waves and wave fronts	277
10.2.2	Sinusoidal plane waves and wavefronts	278
10.2.3	Spherical waves	279
10.2.4	Complex representation of sinusoidal waves	280
10.2.5	Why focus on sinusoidal waves?	281
10.3	Electromagnetic waves in vacuum	281
10.3.1	Propagation equation for the electric field	281
10.3.2	Propagation equation for the magnetic field	282
10.3.3	Propagation in vacuum	282
10.3.4	Structure of electromagnetic plane waves	283
10.3.5	Polarization	284
10.4	Sinusoidal electromagnetic plane waves	286
10.5	Electromagnetic energy and Poynting's theorem	287
10.5.1	Poynting vector and density of electromagnetic energy	288
10.5.2	Poynting vector of a sinusoidal plane wave	289
10.6	Electromagnetic waves in matter	291
10.6.1	General conservation principle	292
10.6.2	Displacement current and Maxwell's equations in matter	293
10.6.3	Poynting's theorem in matter	296
10.6.4	Propagation of waves in dielectric media	297
10.6.5	Resolution of propagation equations in matter	299
10.6.6	Sinusoidal plane waves	301
10.6.7	Propagation in dielectrics	305
10.7	Summary and essential formulas	307
11	Electromagnetic waves in metals and dielectrics, waveguides	311
11.1	Propagation of waves in media with absorption	312
11.1.1	Summary on wave propagation equations and their plane wave solutions	312
11.1.2	Propagation with absorption	314
11.2	Electromagnetic waves in conductors	315
11.2.1	Losses in conductors	316

11.2.2	Conductivity in a time-dependent regime: the Drude-Lorentz model	316
11.2.3	The effective refractive index of metals	317
11.3	Electromagnetic waves in dielectric media	321
11.3.1	The Lorentz model for the classical atom (1878)	321
11.3.2	The complex permittivity	324
11.3.3	The complex refractive index	324
11.4	The laws of reflection and refraction	326
11.4.1	Introduction	326
11.4.2	Continuity equations	326
11.4.3	Reflection and refraction of an electromagnetic wave at an interface	328
11.5	Dielectric waveguides	332
11.6	Fresnel's coefficients	337
11.6.1	Fresnel's coefficients for perpendicular <i>s</i> -polarization	338
11.6.2	Fresnel's coefficients for parallel <i>p</i> -polarization	339
11.6.3	Brewster's angle and polarization by reflection	340
11.6.4	Fresnel's coefficients for an arbitrary polarization	342
11.6.5	Normal incidence	343
11.7	Reflection by metals and metallic waveguides	343
11.8	Reflectance and transmittance	349
11.9	Appendix: Dispersion, temporal response and causality	350
11.9.1	Dispersion	350
11.9.2	Causality	351
11.9.3	Normal dispersion, Cauchy's law	353
11.10	Appendix: Dispersion and propagation of wave packets	354
11.10.1	Wave packets	355
11.10.2	Evolution of a wave packet in the linear regime	356
11.10.3	Dispersion and deformation of wave packets	357
11.10.4	An application of dispersion that led to a Nobel prize	358
11.11	Summary and essential formulas	359
12	Electromagnetic radiation	365
12.1	Retarded Potentials - General Solution to Maxwell's Equations	366
12.1.1	Retarded potentials - Solution of the inhomogeneous wave equation	367
12.2	Radiation	370
12.2.1	Scalar potential	371
12.2.2	Vector potential	372
12.2.3	Radiation fields	372
12.2.4	Poynting's vector and radiated energy	376
12.3	Electric dipole radiation	378
12.3.1	Scalar potential for the electric dipole radiation	379
12.3.2	Current associated with a dipole	380
12.3.3	Vector potential for electric dipole radiation	380
12.3.4	Fields for electric dipole radiation	381

12.3.5	Poynting vector and radiation power for electric dipole radiation	382
12.3.6	Quasi-plane wave approximation	382
12.4	Thin-wire antennas	384
12.5	Scattering	386
12.5.1	Thomson scattering	386
12.5.2	Scattering from an atom or a molecule	387
12.6	Exercises	389
12.7	Summary and essential formulas	402
13	Special relativity	405
13.1	Introduction	406
13.2	The relativity principle	406
13.2.1	Galileo's relativity principle in Newtonian mechanics	406
13.2.2	Electromagnetism and relativity principle	408
13.2.3	Experimental results	410
13.2.4	Einstein's special relativity principle (1905)	413
13.3	Relativistic kinematics	414
13.3.1	Relative nature of time	414
13.3.2	Lorentz's special transformation	416
13.3.3	Invariance of space-time intervals	417
13.3.4	Time dilation	418
13.3.5	Length contraction	421
13.3.6	Relativistic velocity addition	422
13.3.7	Space-time interval between two events	424
13.3.8	Four vectors	426
13.3.9	General Lorentz's transformation	427
13.4	Relativistic dynamics	428
13.4.1	Four-velocity and four-acceleration	428
13.4.2	Relativistic momentum and energy	429
13.4.3	Relation between momentum and energy	430
13.4.4	Relativistic particle dynamics	431
13.4.5	Conservation of power	431
13.5	Summary and essential formulas	432
14	Covariant electrodynamics	433
14.1	Introduction	434
14.1.1	Einstein special relativity principle (1905)	434
14.1.2	Lorentz's transformation	434
14.1.3	Lorentz's transformation in four-vector notation	436
14.2	Electromagnetic quantities	437
14.3	Continuity equation	438
14.4	Lorenz gauge potentials	439
14.5	Field transformation laws	440
14.6	Plane waves	443

14.7	Tensors	444
14.7.1	Cartesian tensors	444
14.7.2	Lorentz tensors	445
14.8	Covariant formulation of Maxwell's equations	445
14.8.1	Inhomogeneous Maxwell's equations	446
14.8.2	Field-field Maxwell's equations	447
14.8.3	Dual tensor	448
14.9	Conservation laws	449
14.10	Summary	450
A	Mathematical toolbox for electrodynamics	451
A.1	Orthogonal coordinates	451
A.2	Vectorial operators	452
A.2.1	Variation of a field	454
A.2.2	Vectorial operators in cylindrical coordinates	455
A.2.3	Vectorial operators in spherical coordinates	456
A.2.4	Vector calculus identities	456
A.2.5	Ostrogradski's theorem	458
A.2.6	Stokes's theorem	459
B	The Dirac distribution	461
B.1	Functions that tend to the Dirac distribution	461
B.2	Solution of Poisson's equation in a finite volume	463
B.2.1	Green's theorem	463
B.2.2	Uniqueness of the solution to the Poisson-Dirichlet problem	464
B.2.3	Uniqueness of the solution to the Poisson-Neumann problem	465
C	From discrete to continuous matter	467
C.1	Spatial averaging	467
C.2	Continuous charge distributions	469
D	The solid angle	471
E	Legendre polynomials	473
F	Spherical harmonics	475

Chapter 1

Maxwell's equations - Electrostatics



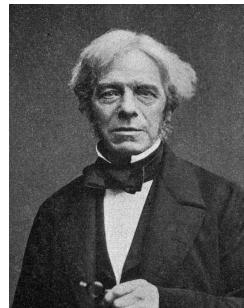
Coulomb
1736–1806



Ampère
1775–1836



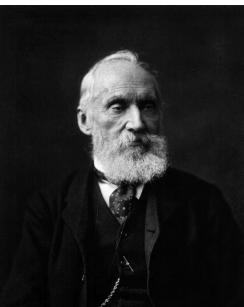
Gauss
1777–1855



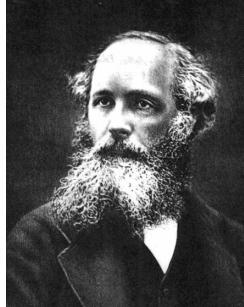
Faraday
1791—1867



Henry
1797–1878



Thomson/Kelvin
1824–1907



Maxwell
1831–1879



Hertz
1857–1894

Introduction

Electrodynamics is the physics of electromagnetic radiation and its interaction with matter. The term *dynamics* refers to the fact that radiation can undergo rapid variations. It is a broad field and one of the fundamental pillars of classical physics, developed from the end of the 18th century. It is based on Maxwell's equations, which successfully describe the behavior of electromagnetic radiation in a huge range (more than 30 orders of magnitude) of the electromagnetic spectrum, including radio waves, microwaves, infrared, visible light, ultraviolet, X-rays, and

gamma rays.

Maxwell's equations were published in 1861 as a unified mathematical model for electric, magnetic, optical and wave phenomena, previously described separately. Maxwell's theory connected electromagnetic waves and light. Together with the Lorentz force that drives the motion of charged particles subjected to an electromagnetic field, Maxwell's equations form the foundation of electromagnetism and light. While Maxwell's theory was proposed to explain accumulated observations of electric and magnetic phenomena during the first half of the 19th century, they serve as postulates for classical electromagnetism.

The goal of this chapter is to provide a reminder of Maxwell's equations taught in the course of *Electromagnetism and light*, and to review some of its basic properties, notably in electrostatics.

1.1 Postulates of electromagnetism

Originally, Maxwell's equation were written in the form of 20 scalar equations in 20 variables, including the force known today as the Lorentz force. An English mathematician and physicist, Oliver Heaviside, developed the vector calculus formalism in 1893 by introducing divergence and curl operators allowing a great reformulation of 12 of the original Maxwell's equations into four equations in the variables \mathbf{E} , \mathbf{B} , ρ and \mathbf{J} , which represent the space and time dependent electric field, magnetic field, charge density and current density, respectively. This reformulation constitutes the *modern* form of Maxwell's equation taught in all textbooks of classical electromagnetism.

There are several ways to associate Maxwell's equations in pairs. (i) As will be seen below, two equations are scalar and involve a divergence operator while the two others are vectorial and involve a curl. (ii) When fields are time independent, Maxwell's equations decouple into a set of two equations involving only the electric field, studied in electrostatics, and a second set involving only the magnetic field, studied in magnetostatics. (iii) Here, we used a third classification and chose to distinguish the pair involving *external* source terms (charge and current densities) from the pair involving only the electric and magnetic fields.

1.1.1 Field-field Maxwell's equations

Field-field Maxwell's equations connect the electric $\mathbf{E}(\mathbf{r}, t)$ and magnetic $\mathbf{B}(\mathbf{r}, t)$ fields. The (\mathbf{E}, \mathbf{B}) pair constitutes the electromagnetic field.

Definition 1.1: Electromagnetic field from potentials

The electromagnetic field at point \mathbf{r} , time t , is entirely described by the knowledge of

- a vector potential $\mathbf{A}(\mathbf{r}, t)$,
- a scalar potential $V(\mathbf{r}, t)$.

The electric and magnetic fields are then defined from the potentials by the relations

$$\boxed{\begin{aligned}\mathbf{E}(\mathbf{r}, t) &= -\nabla V(\mathbf{r}, t) - \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}, \\ \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t).\end{aligned}}$$

From the electromagnetic field definition, taking the curl of the electric field, we have

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\nabla \times \nabla V(\mathbf{r}, t) - \nabla \times \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}.$$

The curl of a gradient is zero, so the first term on the right-hand-side is zero. In the second term on the right-hand-side, time and space are independent variables, so the curl and the time derivative can be swapped, resulting in the time derivative of $\nabla \times \mathbf{A}$. The latter quantity is \mathbf{B} by definition¹. Thus,

$$\nabla \times \mathbf{E} = -\frac{\partial(\nabla \times \mathbf{A})}{\partial t} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Hence, we find Faraday-Maxwell's equation stating that a variation in time of the magnetic field is responsible for the generation of an electric field. Recall that Faraday-Maxwell's equation governs the phenomenon of induction.

Taking now the divergence of the magnetic field, we find zero because the divergence of a curl is zero.

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}(\mathbf{r}, t)) = 0.$$

Hence we find Thomson-Maxwell's equation stating that there is no magnetic monopole and that the flux of the magnetic field through a cross-section of a field tube is preserved. Summarizing Field-field Maxwell's equations,

Definition 1.2: Field-field Maxwell's equations

The pair of source-field Maxwell's equations read

$$\boxed{\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}, t) &= -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t}, && \text{Faraday-Maxwell} \\ \nabla \cdot \mathbf{B}(\mathbf{r}, t) &= 0. && \text{Thomson-Maxwell}\end{aligned}}$$

Note that Faraday-Maxwell's and Thomson-Maxwell's equation follow directly from the definition of the fields \mathbf{E}, \mathbf{B} from the potentials V, \mathbf{A} . We may as well consider Faraday-Maxwell's and Thomson-Maxwell's equation as postulates. In this case:

- vectorial analysis, tells us that if a field is divergence-free, it can be written as the curl of a vector field. Hence, if $\nabla \cdot \mathbf{B} = 0$, a vector potential \mathbf{A} exists such that $\mathbf{B} = \nabla \times \mathbf{A}$.

¹Space and time dependence of all fields, valid up to section 1.4, will be omitted to facilitate readability.

- vectorial analysis tells us that if a field is curl-free, it can be written as the gradient of a scalar potential. Hence, if Faraday-Maxwell's equation is satisfied by \mathbf{E} and $\mathbf{B} = \nabla \times \mathbf{A}$, then $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$ is curl free, which means it can be written as $-\nabla V$, which is the definition of the electric field $\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}$.

Note that the potentials are not unique. Indeed, consider the vector potential $\mathbf{A}' = \mathbf{A} + \nabla \phi$ and the scalar potential $V' = V - \frac{\partial \phi}{\partial t}$, where ϕ is any scalar function depending on space and time. We have

$$\mathbf{B}' = \nabla \times \mathbf{A}' = \nabla \times \mathbf{A} + \nabla \times \nabla \phi = \nabla \times \mathbf{A} = \mathbf{B},$$

and

$$\mathbf{E}' = -\nabla V' - \frac{\partial \mathbf{A}'}{\partial t} = -\nabla V + \nabla \frac{\partial \phi}{\partial t} - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial \nabla \phi}{\partial t} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E}.$$

Therefore, \mathbf{A}' and V' define the same electromagnetic field as \mathbf{A} and V .

In order for the potentials to be uniquely defined, we must in general require them to satisfy an additional constraint, called a gauge condition.

1.1.2 Field-source Maxwell's equations

Field-source Maxwell's equations connect sources for the electromagnetic field to the fields. Let us consider *vacuum* in which we allow for the presence of charges, either at rest or in motion. In this chapter, we exclude, however, *bound* charges present in neutral matter, which will be treated in chapter 4. The sources for the electromagnetic field are the charges at rest and in motion. Let us assume that we can describe the charges by

- a volume charge density $\rho(\mathbf{r}, t)$ (unit: C m^{-3})
- a current density vector $\mathbf{J}(\mathbf{r}, t)$ (unit: A m^{-2})

These quantities refer to continuous charge distributions. However we will see below in section 1.4.3 that all quantities appearing in Maxwell's equations are spatially averaged over a *mesoscopic* volume². Consequently even discrete charge distributions can be formally assigned a charge density and current density. Consider for instance point charges q_i of velocity \mathbf{v}_i . The associated charge and current densities read

$$\begin{aligned}\rho(\mathbf{r}, t) &= \sum_{i \in \mathcal{V}(\mathbf{r})} q_i \phi(\mathbf{r}) \\ J(\mathbf{r}, t) &= \sum_{i \in \mathcal{V}(\mathbf{r})} q_i \mathbf{v}_i \phi(\mathbf{r})\end{aligned}$$

²A mesoscopic volume is very large compared to the typical size of an atom, and thus contains a very large number of particles (atoms, charges, etc). Yet, it is very small compared to the typical macroscopic dimension of a charge distribution, allowing us to spatially average fields as well as charge and current densities over a mesoscopic volume around each point of space, so as to obtain local values that are not singular when approaching a *point* charge.

where the averaging function $\phi(\mathbf{r})$ is equal to a constant $1/\mathcal{V}$ in a volume \mathcal{V} centered at \mathbf{r} and 0 otherwise. All charges within a small volume \mathcal{V} around every point thus contribute to the average charge density and current density at that point and since charges are in motion, they may enter and leave the volume around a given point, leading to the time dependence of the average charge and current densities. This spatial averaging procedure allows us to define an average density $n(\mathbf{r}, t)$ of mobile charges and an average velocity of charges $\mathbf{v}(\mathbf{r}, t)$, that is, at any time t , the charges within a small volume \mathcal{V} around each point \mathbf{r} form a *fluid* of electricity characterized by its local density $n(\mathbf{r}, t)$ and local velocity $\mathbf{v}(\mathbf{r}, t)$.

Definition 1.3: Current density

Let n be the density of mobile charges (each of charge q) in a distribution and \mathbf{v} the average velocity of these charge carriers. The current density is defined as

$$\mathbf{J} = qn\mathbf{v} \quad (1.1)$$

whose unit in the SI system is $\text{Cs}^{-2} \text{m}^{-1} = \text{A}^2 \text{m}^{-1}$.

Coming back to Maxwell's equations,

Definition 1.4: Source-field Maxwell's equations

The pair of source-field Maxwell's equations read

$\nabla \cdot (\epsilon_0 \mathbf{E})$	$= \rho(\mathbf{r}, t)$	Gauss-Maxwell
$\nabla \times \left(\frac{\mathbf{B}}{\mu_0} \right)$	$= \mathbf{j}(\mathbf{r}, t) + \underbrace{\frac{\partial(\epsilon_0 \mathbf{E})}{\partial t}}_{\text{displacement current}}$	Ampère-Maxwell

The second term on the right-hand-side of Ampère-Maxwell's equation is called the displacement current. It allowed Maxwell to predict the propagation of waves (radio-waves, light) through empty space, a phenomenon observed for radio-waves by Hertz in 1886, thereby providing experimental confirmation of Maxwell's theory.

- **REMARKS**

- The symbol μ_0 denotes the vacuum magnetic permeability. Its value and units are

$$\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1} \text{ (Henry per meter)},$$

where the Henry is the SI unit used to measure inductance coefficients (see chapter 8):

$$1 \text{ H} = 1 \text{ m}^2 \text{ kg s}^{-2} \text{ A}^{-2}.$$

- The symbol ϵ_0 denotes the dielectric permittivity of vacuum. Its relation to μ_0 is given by the relation $\epsilon_0\mu_0c^2 = 1$, where $c \sim 3 \times 10^8 \text{ m s}^{-1}$ is the light velocity in vacuum. Therefore, ϵ_0 takes the value

$$\epsilon_0 \sim \frac{1}{36\pi \times 10^9} \sim 8.85 \times 10^{-12} \text{ F m}^{-1} \text{ (Farad per meter)},$$

where the Farad is the S.I. unit used to measure capacitance (see chapter 3):

$$1 \text{ F} = 1 \text{ m}^2 \text{ kg}^{-1} \text{ s}^4 \text{ A}^2.$$

Taking the divergence of Ampère-Maxwell's equation results in the charge conservation equation

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0$$

charge conservation.

Indeed, the divergence of a curl is zero, so we only need to consider the divergence of the right hand-side of Ampère-Maxwell's equation. $\nabla \cdot \mathbf{J}$ is obtained from the first term on the right-hand-side. The divergence operator and the differentiation with respect to time can be swapped (independent time and space variables). The second term on the right hand side is then expressed, using Gauss-Maxwell's equation, as $\frac{\partial}{\partial t} \nabla \cdot (\epsilon_0 \mathbf{E}) = \frac{\partial}{\partial t} \rho$, which completes the derivation of the charge conservation equation.

1.1.3 The Lorentz force

The Lorentz force describes the action of an electromagnetic field on a charged particle.

Consider a particle of mass m and charge q . A principle of classical mechanics states that the mass is invariant with respect to a change of inertial reference frame. The charge follows the same property and this is still true in relativity. Let $\mathbf{r} \equiv (x, y, z)$ denote the position of the particle, and \mathbf{v} , its velocity.

The Lorentz force on a charge in an inertial reference frame writes

$$\mathbf{F}_L = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}.$$

The motion of the charge driven by the Lorentz force is governed by an equation of motion. For instance, for a nonrelativistic particle, Newton's equation of classical mechanics allows us to know the position and velocity of the particle.

1.2 Gauss's theorem in vacuum

Recall Gauss's law already seen in electrostatics: it states that the electric flux through a closed surface is proportional to the charge enclosed by it.

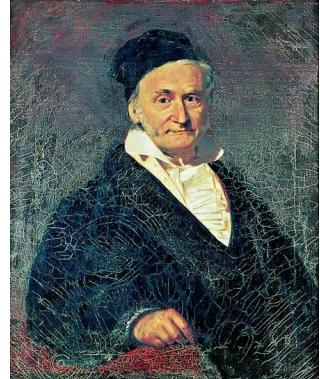
Definition 1.5: Gauss's theorem

For every closed surface $\partial\Omega$, the flux of the electric field through $\partial\Omega$ is proportional to $Q_\Omega(t)$, the total charge enclosed by $\partial\Omega$:

$$\Phi_{\partial\Omega, \mathbf{E}} = \iint_{\partial\Omega} \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r}) = \frac{Q_\Omega(t)}{\epsilon_0} \quad (1.2)$$

Writing explicitly the charge enclosed by $\partial\Omega$ as an integral of the charge density ρ over Ω , Gauss's law is rewritten as

$$\iint_{\partial\Omega} \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r}) = \frac{1}{\epsilon_0} \iiint_{\Omega} \rho(\mathbf{r}) d^3 r.$$

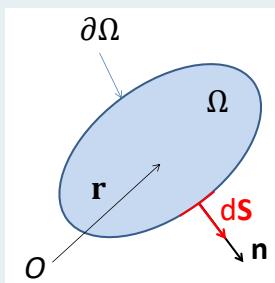


Carl Friedrich GAUSS
(1777-1855), German mathematician and physicist.

Gauss's law is very useful for calculating the electric field of charge distributions having spatial symmetries. Note also that electric fields created by charges outside of Ω do not contribute to the total flux.

The Gauss theorem follows directly from Gauss-Maxwell's equation. As stated by the postulates of electromagnetism, it is valid in general in the time-dependent regime, not only in the context of electrostatics. From Gauss-Maxwell's equation to the Gauss theorem, a convenient way is to use Green-Ostrogradsky's theorem.

Definition 1.6: Green-Ostrogradski's (flux-divergence) theorem



For every closed surface $\partial\Omega$, the flux of a vectorfield \mathbf{F} through $\partial\Omega$ is equal to the volume integral of $\nabla \cdot \mathbf{F}(\mathbf{r})$, the divergence of \mathbf{F} ,

$$\iint_{\partial\Omega} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{S}(\mathbf{r}) = \iiint_{\Omega} \nabla \cdot \mathbf{F}(\mathbf{r}) d^3 r.$$

The flux-divergence theorem applied to the field $\mathbf{F} \equiv \epsilon_0 \mathbf{E}(\mathbf{r}, t)$ almost directly leads to Gauss's theorem. Indeed,

$$\oint\!\!\!\oint_{\partial\Omega} \epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{n} dS = \iiint_{\Omega} \nabla \cdot (\epsilon_0 \mathbf{E}(\mathbf{r}, t)) d^3r$$

Gauss-Maxwell's equation states $\nabla \cdot \epsilon_0 \mathbf{E}(\mathbf{r}, t) = \rho(\mathbf{r}, t)$, hence,

$$\oint\!\!\!\oint_{\partial\Omega} \epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{n} dS = \iiint_{\Omega} \rho(\mathbf{r}, t) d^3r,$$

where we recognize on the right hand side the total charge enclosed in the volume Ω , $Q(t) = \iiint_{\Omega} \rho(\mathbf{r}, t) d^3r$, which is possibly time dependent. Gauss's theorem

$$\oint\!\!\!\oint_{\partial\Omega} \epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{n}(\mathbf{r}) dS = Q(t)$$

is nothing but the integral form of Gauss-Maxwell's equation. Both are valid in general, even if the charge density and the electric field are time dependent quantities.

1.3 Charge conservation

We have seen that the charge conservation equation written in local form

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0 \quad (1.3)$$

can be derived from Maxwell's equations, that is, charge conservation is a consequence of the postulates of electromagnetism.

In this section, we will establish an integral form of the charge conservation equation for a macroscopic volume Ω , without assuming (1.3). In this aim, we will apply a very general principle of conservation reading

$$\text{Variation within } \Omega = -(\text{Flux through } \partial\Omega) + \text{creation} - \text{losses}$$

This principle simply states that the variation or time rate of change of a quantity Q contained in a volume Ω is equal to the sum of the quantity crossing the boundary $\partial\Omega$ (the minus sign indicates losses as the outward flux is positive) with the quantities created or destroyed (losses) within Ω , for instance by a chemical reaction. In our case, this general principle is applied to the total charge contained in a volume Ω .

$$Q(t) = \iiint_{\Omega} \rho(\mathbf{r}, t) d^3r$$

The relation (1.3) between the charge density ρ and the current density \mathbf{J} will then be shown to result from the integral form of the charge conservation equation, which ensures the local conservation of charge at every point and at any time.

1.3.1 Current through a surface

The flux of the current density \mathbf{J} through a surface represents the amount of charge per unit time going through that surface. Indeed, consider an arbitrary infinitesimal surface dS with normal \mathbf{n} .

During a time interval dt , all the charges that are contained inside a volume $d\mathbf{v} \cdot \mathbf{n} dS$ have gone through dS . Therefore, the charge dq that goes through dS during the time interval dt is

$$dq = qndS\mathbf{v} \cdot \mathbf{n} dt = \underbrace{qn}_{\mathbf{J}} \mathbf{v} \cdot d\mathbf{S} dt$$

and the charge going through dS per unit time is

$$\frac{dq}{dt} = \mathbf{J} \cdot \mathbf{n} dS = \Phi_{dS, \mathbf{J}}.$$

Note that for a current density of electrons, $qn < 0$ so that the average speed \mathbf{v} and the associated current density \mathbf{J} have directions of opposite sign. It is purely a matter of convention then that a positive flux of electrons through a surface corresponds to a negative current.

Definition 1.7: Electric current

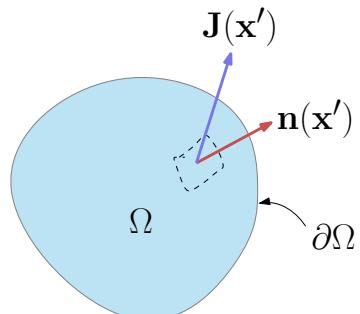
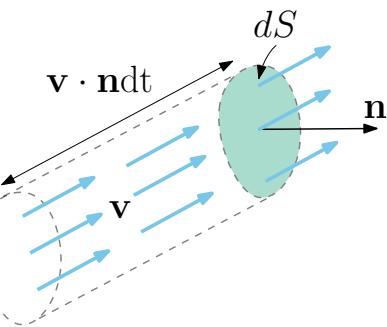
The current I through a surface S of a conductor is defined as the total amount of charge going through S per unit time. According to the definition of the current density \mathbf{J} , this can be written as the flux of \mathbf{J} through the surface S

$$I = \frac{dq}{dt} = \iint_S dS(\mathbf{x}') \mathbf{n}(\mathbf{x}') \cdot \mathbf{J}(\mathbf{x}') = \Phi_{S, \mathbf{J}} \quad (1.4)$$

The unit for the electric current is the Ampère (A), with $1 \text{ A} = 1 \text{ C s}^{-1}$.

1.3.2 Charge conservation law - Integral form

Let us consider a volume Ω and its boundary $\partial\Omega$. We also assume that there is no source or sink of charges within Ω , that is no process that creates or destroy charges can occur within Ω . Charges can only cross the boundary $\partial\Omega$ and increase $dQ/dt > 0$ or decrease $dQ/dt < 0$ the total charge $Q(t)$ contained in Ω .



At time t , the amount of charge going through the surface $\partial\Omega$ per unit time is the current given by

$$I(t) = \Phi_{\partial\Omega, \mathbf{J}(t)} = \iint_{\partial\Omega} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r}).$$

Since charge is conserved, a positive (resp. negative) $I(t)$ implies that there must be a decrease (resp. increase) of the total charge $Q(t)$ contained in Ω . That is $dQ(t)/dt = -I(t)$ or, equivalently:

$$\frac{dQ}{dt} = \underbrace{\frac{d}{dt} \iiint_{\Omega} \rho(\mathbf{r}, t) d^3 r}_{\text{Variation}} = - \underbrace{\iint_{\partial\Omega} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r})}_{\text{Flux}}$$

which is the integral form of the charge conservation law in the general form

$$\text{Variation within } \Omega = -\text{Flux through } \partial\Omega.$$

1.3.3 Charge conservation law - Differential form

The time derivative can be swapped with the integral over space variables in the integral form of charge conservation, which can then be rewritten as

$$\iiint_{\Omega} \frac{\partial \rho(\mathbf{r}', t)}{\partial t} d^3 r' + \iint_{\partial\Omega} d\mathbf{S}(\mathbf{r}') \cdot \mathbf{J}(\mathbf{r}', t) = 0. \quad (1.5)$$

Using Green-Ostrogradsky's theorem, we have

$$\iint_{\partial\Omega} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r}) = \iiint_{\Omega} \nabla \cdot \mathbf{J}(\mathbf{r}, t) d^3 r,$$

and inserting in (1.5), we find

$$\iiint_{\Omega} \left(\nabla' \cdot \mathbf{J}(\mathbf{r}', t) + \frac{\partial \rho(\mathbf{r}', t)}{\partial t} \right) d^3 r' = 0 \quad \forall \Omega$$

from which we obtain the differential form of the charge conservation law

$$\nabla \cdot \mathbf{J}(\mathbf{r}, t) + \frac{\partial \rho(\mathbf{r}, t)}{\partial t} = 0 \quad (1.6)$$

From this, we see that a current density compatible with a steady state regime ($\partial\rho/\partial t = 0$) is such that

$$\nabla \cdot \mathbf{J}(\mathbf{r}) = 0 \quad (1.7)$$

which means that the total current across any closed surface must be zero in the steady state.

1.4 The electrostatic field

In this section, we consider the time independent regime, for which none of the quantities that appear in Maxwell's equations depend on time. In this case, all time derivatives in Maxwell's equations vanishes. We observe that the four Maxwell equations are decoupled into two pairs, involving only \mathbf{E} or only \mathbf{B} . Focusing on the pair involving the electric field, we have

$$\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}) &= \mathbf{0}, \\ \nabla \cdot \epsilon_0 \mathbf{E}(\mathbf{r}) &= \rho(\mathbf{r}),\end{aligned}$$

where the first equation $\nabla \times \mathbf{E}(\mathbf{r}) = 0$, is equivalent to say that there exist a scalar potential $V(\mathbf{r})$ such that

$$\mathbf{E}(\mathbf{r}) = -\nabla V(\mathbf{r})$$

This is the regime of electrostatics.

1.4.1 The Laplace and Poisson equations

Combining $\mathbf{E}(\mathbf{r}) = -\nabla V(\mathbf{r})$ with Gauss-Maxwell's equation, we obtain:

$$\nabla \cdot \underbrace{\mathbf{E}(\mathbf{r})}_{-\nabla V} = -\nabla^2 V(\mathbf{r}) = \frac{\rho(\mathbf{r})}{\epsilon_0}$$

where $\nabla^2 = \nabla \cdot \nabla$ is the Laplace operator. In Cartesian coordinates,

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \frac{-\rho(x, y, z)}{\epsilon_0}$$

Definition 1.8: Poisson's equation (1813)

The electric potential V satisfies Poisson's equation

$$\nabla^2 V(\mathbf{x}) = -\frac{\rho(\mathbf{x})}{\epsilon_0} \tag{1.8}$$

which, in regions free of charges, reduces to Laplace's equation

$$\nabla^2 V(\mathbf{r}) = 0 \tag{1.9}$$



Left: Simeon Denis POISSON (1781-1840), French mathematician and physicist, professor of mathematics (1806-1815) at Ecole Polytechnique. Right: Pierre-Simon LAPLACE (1749-1827), French mathematician and physicist.

1.4.2 Boundary conditions for Poisson's equation

Often, we are only interested in determining the solution of Poisson's equation inside a finite volume $\Omega \subset \mathbb{R}^3$. Mathematically, it can be shown that all of the information outside Ω can be translated into a proper boundary condition for the potential in $\partial\Omega$. Two common boundary conditions, for which the solution of Poisson's equation is unique, are the following:

- Dirichlet's boundary condition: one specifies the value of the potential V at the boundary $\partial\Omega$

$$\mathbf{r} \in \partial\Omega : V(\mathbf{r}) = V_D(\mathbf{r})$$

- Neumann's boundary condition: one specifies the normal component of the electric field $\mathbf{E}(\mathbf{r})$ (or potential gradient) at the boundary $\partial\Omega$

$$\mathbf{r} \in \partial\Omega : \mathbf{E}(\mathbf{r}) \cdot \mathbf{n} = E_N(\mathbf{r})$$

The reader can refer to the appendix equation xxx for a proper mathematical solution of Poisson's equation in a finite volume.

1.4.3 Solutions to the Poisson equation

• COULOMB'S POTENTIAL AND FIELD

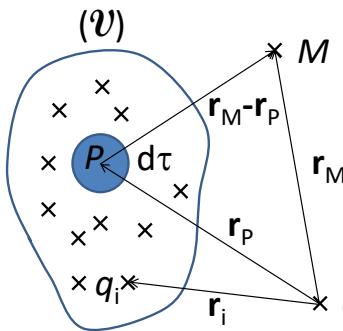
Recall that the potential and electric field generated at any point \mathbf{r} of space by a point charge q located at \mathbf{r}_0 reads

$$V(\mathbf{r}) = \frac{q}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}_0|}, \quad \text{and} \quad \mathbf{E}(\mathbf{r}) = \frac{q(\mathbf{r} - \mathbf{r}_0)}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}_0|^3}.$$

Obviously, these expressions satisfy Laplace's equation $\Delta V(\mathbf{r}) = 0 \forall \mathbf{r} \neq \mathbf{r}_0$.

• DISCRETE DISTRIBUTION OF CHARGES

Let us assume we have now a set of discrete point charges q_i located at \mathbf{r}_i . The superposition principle allows us to write the potential generated by these charges at any point \mathbf{r} as the sum of Coulomb's laws associated to each charge,



$$V(\mathbf{r}) = \sum_i \frac{q_i}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}_i|}$$

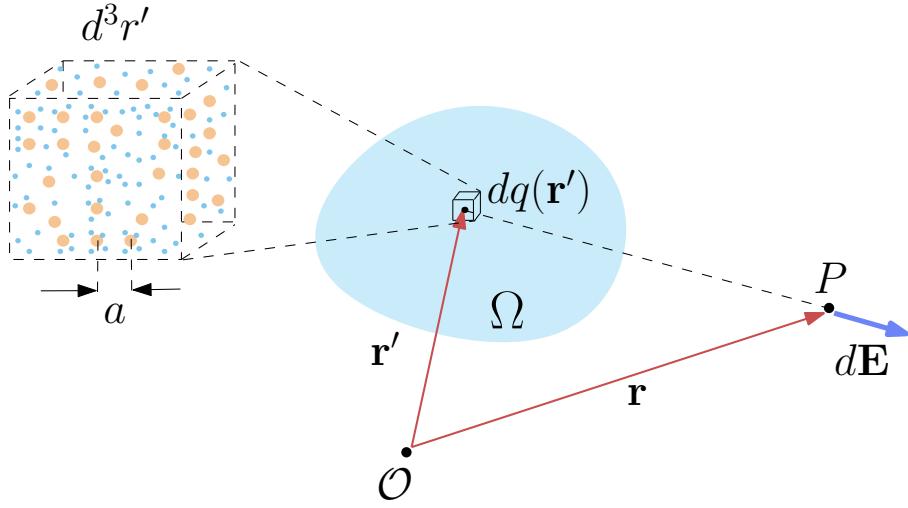
Similarly, we write the electric field generated by the set of point charges as

$$\mathbf{E}(\mathbf{r}) = \sum_i \frac{q_i(\mathbf{r} - \mathbf{r}_i)}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}_i|^3}.$$

• FROM DISCRETE TO CONTINUOUS CHARGE DISTRIBUTIONS

Any piece of matter is made of elementary particles, and so described by a discrete set of point charges separated by vacuum. In a solid, for example, neighboring atoms are separated by a distance of the order of $a \sim 10^{-10}$ m, so that the electric field exhibit rapid variations over distances of the order of a . However, when dealing with macroscopic objects, we are not capable of resolving such rapid variations at the atomic scale, and in consequence we are only interested in averaged values of both the charge spatial distribution and the resulting electric field. The idea is therefore to replace a discrete set of charges by a continuous medium in which a smooth charge density can be defined at every point.

Let Ω be the volume occupied by a macroscopic object. At any point $\mathbf{r}' \in \Omega$, consider a sub-volume d^3r' around \mathbf{r} such that $a^3 \ll d^3r' \ll \Omega$, containing thus a very large number N of point charges. Within d^3r' , there is a total charge $dq(\mathbf{r}') = \sum_{i=1}^N q_i$.



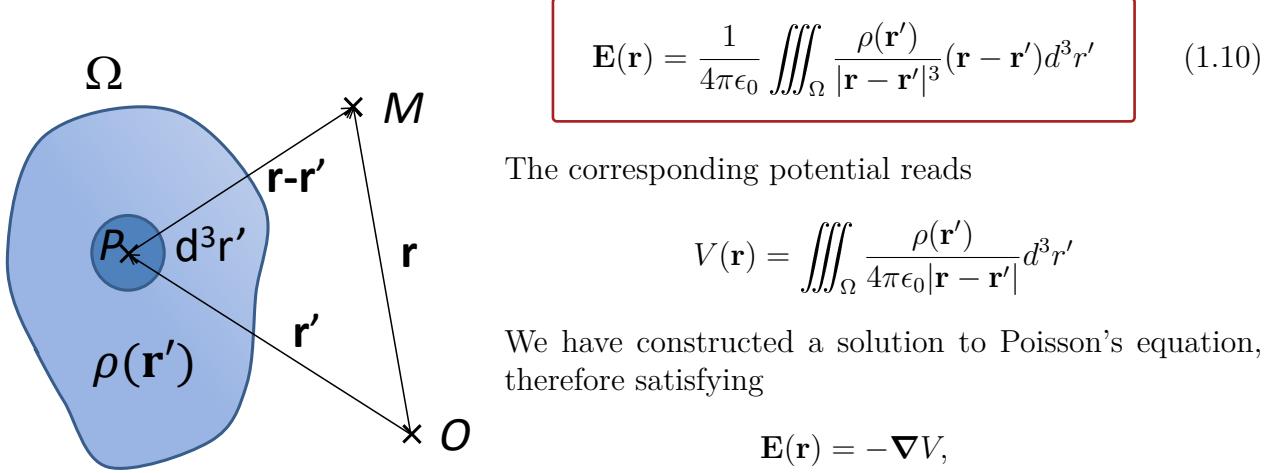
The idea is to consider d^3r' as being a homogeneously charged volume, that is, we associate to \mathbf{r}' a certain average charge density per unit volume $\rho(\mathbf{r}')$ (C/m^3) such that

$$dq(\mathbf{r}') = \rho(\mathbf{r}')d^3r'$$

At this stage, the real charge density has been replaced by its spatial average over a length scale much larger than a . The field generated at a point \mathbf{r} by $dq(\mathbf{r}')$ will be approximately that of a point charge:

$$d\mathbf{E}(\mathbf{r}) = \frac{\rho(\mathbf{r}')d^3r'}{4\pi\epsilon_0} \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}$$

To obtain the field generated by the whole charge distribution, it is enough, according to the superposition principle, to add the individual contributions of each elementary volume within Ω . We therefore sum over the possible values of $\mathbf{r}' \in \Omega$ and the following integral formula is obtained:



When one of the dimensions of Ω is much smaller than the other ones (for example, when the charge is concentrated over a very small thickness), one may go further and consider a two-

dimensional charge distribution over a surface S , so that to each $\mathbf{r}' \in S$ we associate a surface charge density, denoted σ to distinguish it from a volume charge density, and

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iint_S \frac{\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} (\mathbf{r} - \mathbf{r}') dS'$$

Finally, for a one-dimensional charge distribution over a curve Γ , we consider a line charge density, typically denoted λ , so that the field at \mathbf{r} is given by

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\Gamma} \frac{\lambda(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} (\mathbf{r} - \mathbf{r}') dl'$$

A more rigorous justification of this method of spatial averaging is given in the appendix C.

1.4.4 The Dirac distribution, the mathematical problem of dealing with point charges

As seen in section 1.4.3, performing a spatial average of discrete charge distributions has allowed us to model a charged object Ω by a function $\rho : \Omega \rightarrow \mathbb{R}$ which is continuous inside Ω and that represents the averaged charge density per unit volume at all points of Ω . This density allows the electric field to be calculated in all space by means of the integral (1.10). Let us know reverse the problem: can we express a point charge via a charge density ρ ? If yes, this would mean that we can consider (1.10) as a completely general expression of the electric field, whether it is generated by a discrete set of charges or a continuous charge distribution.

Indeed, in physics it is frequent to resort to the notion of a point object to deal, for example, with fundamental particles such as the electron or with objects whose dimensions are negligible with respect to the relevant length scales of the problem. Here, our objective is to assign to a charge q located at point \mathbf{r}_0 a charge density ρ . Assuming that this charge density exists, it must vanish everywhere except at x_0 , that is $\rho(\mathbf{r}, \mathbf{r}_0) = 0$ if $\mathbf{r} \neq \mathbf{r}_0$, and its integral over the whole space must equal q , $\iiint_{\mathbb{R}^3} \rho(\mathbf{r}, \mathbf{r}_0) d^3 r = q$. In addition, by replacing ρ in the integral (1.10), we should obtain the electric field of a point charge:

$$\frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \rho(\mathbf{r}, \mathbf{r}_0) \frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} d^3 r' = \frac{q}{4\pi\epsilon_0} \frac{(\mathbf{r} - \mathbf{r}_0)}{|\mathbf{r} - \mathbf{r}_0|^3}$$

We see that the role of $\rho(\mathbf{r}, \mathbf{r}_0)$ in this integral is tantamount to extracting the value of the integrand at $\mathbf{r} = \mathbf{r}_0$. The above required properties can be summarized by writing $\rho(\mathbf{r}, \mathbf{r}_0) = q\delta(\mathbf{r} - \mathbf{r}_0)$, where δ is called the Dirac function, with

$$\delta(\mathbf{r}) = 0 \quad \forall \mathbf{r} \neq \mathbf{0} \tag{1.11}$$

and such that for every regular function φ :

$$\iiint_{\mathbb{R}^3} \delta(\mathbf{r}) \varphi(\mathbf{r}) d^3 r = \varphi(\mathbf{0}). \tag{1.12}$$

In particular, applying (1.12) to the function that is equal to 1 everywhere, we have

$$\iiint_{\mathbb{R}^3} \delta(\mathbf{r}) d^3 r = 1. \quad (1.13)$$

In one dimension, this property writes

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1, \quad (1.14)$$

which indicates that we are looking for a function with the dimension of a volume density in \mathbb{R} or the inverse of a length in one dimension. However, such a δ function does not exist. Indeed, if a function is zero except at a single point, its integral is necessarily zero, in contradiction with Eq. (1.13). The property (1.12) is then impossible to fulfill except for $\varphi(\mathbf{0}) = 0$. In fact, the suitable mathematical description to deal with such point objects is not a function but a *distribution*, whose formalism was properly established by the French mathematician Laurent Schwartz (1915-2002). While the theory of distributions and the proper definition of a Dirac distribution will not be treated in this course, it can be shown that one can *approximate* the δ by a sequence of functions $(\delta_n)_{n \in \mathbb{N}}$, such that $\lim_{n \rightarrow \infty} \delta_n(\mathbf{r}) = 0$ for $\mathbf{r} \neq \mathbf{0}$ and

$$\lim_{n \rightarrow \infty} \iiint_{\mathbb{R}^3} \delta_n(\mathbf{r}) \varphi(\mathbf{r}) d^3 r = \varphi(\mathbf{0}). \quad (1.15)$$

This justify the incorrect notation, used in most textbooks, consisting in assuming the existence of a function δ with the properties (1.11), (1.12) and (1.13).

Let us give examples of functions that can be seen as approximations of the Dirac distribution in one dimension, that is, satisfying $\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x)$: consider a peaked function at zero, whose width $\epsilon_n \equiv 1/n$ tends to zero when $n \rightarrow \infty$ while its peak value tends to ∞ , with the constraint (1.14). For instance, we can construct rectangular functions, or Gaussian functions:

$$\delta_n(x) = \frac{1}{\epsilon_n} \text{rect} \left[\frac{x}{\epsilon_n} \right] \quad \begin{array}{c} \text{Diagram of a rectangle of height } \frac{1}{\epsilon_n} \text{ and width } \epsilon_n \end{array}$$

$$\delta_n(x) = \frac{1}{\epsilon_n \sqrt{\pi}} \exp \left[-\frac{x^2}{\epsilon_n^2} \right] \quad \begin{array}{c} \text{Diagram of a Gaussian curve peaking at } x=0 \text{ with width } \epsilon_n \end{array}$$

By means of a translation of the Dirac distribution, it is easy to use property (1.12) to extract the value of any regular function $\phi(x)$ at any point x_0 . This writes

$$\boxed{\int_{-\infty}^{+\infty} \delta(x - x_0) \phi(x) dx = \phi(x_0)}.$$

Since a Dirac delta makes sense only under an integral, in two or three dimensions, we write

$$\delta^{(3)}(\mathbf{r} - \mathbf{r}_0) = \delta(x - x_0) \delta(y - y_0) \delta(z - z_0),$$

$\delta^{(3)}(\mathbf{r} - \mathbf{r}_0)$ has units of a volume density, where the superscript (3) indicates three dimensions in contrast with the one-dimensional Dirac deltas, with units of the inverse of a length, that appear

on the right-hand-side. This superscript is nevertheless needless as the number of dimensions is usually clear from the context.

Finally, we write here a fundamental identity of electrostatics involving the Dirac distribution, with the incorrect (nevertheless popular) notation:

$$\nabla^2 \left(\frac{1}{|\mathbf{r}|} \right) = -4\pi\delta(\mathbf{r}), \quad (1.16)$$

which is useful to check that the Coulomb's potential is solution Poisson's equation for a point charge.

A brief discussion of the Delta distribution and its main properties are summarized in the appendix xxx

1.4.5 Charge and current densities for discrete charge distributions

All the ingredients are now ready to systematically express the source terms in Maxwell's equations as continuous charge and current densities even in the case of a discrete charge distribution, say q_i at positions $\mathbf{r}_i(t)$, with velocities \mathbf{v}_i . Indeed the distribution can be described by

- a volume density of charges [C/m^3]

$$\rho(\mathbf{r}, t) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i(t))$$

- a volume density of current [A/m^2]

$$\mathbf{j}(\mathbf{r}, t) = \sum_i q_i \mathbf{v}_i(t) \delta(\mathbf{r} - \mathbf{r}_i(t))$$

As an example of good practice of the Dirac distribution, recall that the above densities using $\delta(\cdot)$ make sense only under an integral. For instance, to calculate the total charge within a volume Ω , one would write

$$Q(t) = \iiint_{\Omega} \rho(\mathbf{r}, t) d^3 r = \iiint_{\Omega} \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i(t)) d^3 r = \sum_{i \in \Omega} q_i \underbrace{\iiint_{\Omega} \delta(\mathbf{r} - \mathbf{r}_i(t)) d^3 r}_{1} = \sum_{i \in \Omega} q_i.$$

As a second example, consider the case of electrostatics, for which the potential can be similarly calculated from the general solution to the Poisson equation

$$V(\mathbf{r}) = \iiint_{\Omega} \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} d^3 r'$$

Introducing the pseudo-continuous charge density, we have

$$V(\mathbf{r}) = \iiint_{\Omega} \frac{\sum_{i \in \Omega} q_i \delta(\mathbf{r}' - \mathbf{r}_i)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} d^3 r' = \frac{1}{4\pi\epsilon_0} \sum_{i \in \Omega} q_i \underbrace{\iiint_{\Omega} \frac{\delta(\mathbf{r}' - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}'|} d^3 r'}_{1/|\mathbf{r} - \mathbf{r}_i|}.$$

The δ function is peaked at $\mathbf{r}' = \mathbf{r}_i$, so we extract the value or the regular function under the integral at $\mathbf{r}' = \mathbf{r}_i$ and obtain

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i \in \Omega} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}.$$

We retrieved the Coulomb potential for a discrete charge distribution.

As a final example, let us check that Coulomb's potential for a point charge q at $\mathbf{r}_0 = \mathbf{0}$ satisfies Poisson's equation. Remark first that simply applying the Laplacian³ to a Coulomb potential $V(r) \propto \frac{1}{r}$ yields $\Delta V(r) = 0$ for every $\mathbf{r} \neq \mathbf{0}$. How to deal with the point charge on the right hand side of Poisson's equation?

We can write the charge density in terms of the Dirac distribution as $\rho(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$. We also know that the potential generated by this charge is simply $V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_0|}$. Poisson's equation then writes

$$\nabla^2 V = \frac{1}{4\pi\epsilon_0} \nabla^2 \frac{1}{|\mathbf{r} - \mathbf{r}_0|} = -\frac{\delta(\mathbf{r} - \mathbf{r}_0)}{\epsilon_0}$$

and because of the fundamental equation (1.16), $\nabla^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -4\pi\delta(\mathbf{r} - \mathbf{r}')$, valid in the framework of distributions, we conclude that the Coulomb potential satisfies Poisson's equation

1.5 Symmetry arguments in electrostatics

In practice, Gauss's law is useful whenever the flux integral in 1.2 can be simplified thanks to the existence of spatial symmetries of the electric field and the use of a properly chosen closed surface. In order to determine whether the electric field generated by a charge distribution ρ has spatial symmetries, it is sufficient to study the symmetry properties of ρ , as stated by Curie's principle.

Definition 1.9: Curie's principle (1894)

When some causes (here, a charge density ρ) produce some effects (here, an electric field \mathbf{E}), the symmetries of the causes must also hold for the effects.

Note that the reciprocal is not true, that is, the effects may have more symmetries than their causes.

³in spherical coordinates, that is, for a regular function, $f(r)$, $\Delta f(r) \equiv \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) f(r)$.

As will be shown below, the study of the symmetries of the charge distribution will allow us to:

1. Reduce the number of variables required to describe the spatial variations of \mathbf{E} .
2. Constrain the orientation of the field \mathbf{E} .

1.5.1 Invariance by spatial translation along an axis

Consider the case of a charge distribution that is invariant under translation along an axis (Oz for example). This means, in Cartesian coordinates,

$$\rho(x, y, z + a) = \rho(x, y, z) \quad \forall a \in \mathbb{R}$$

so that the charge density is independent of the z -coordinate, $\rho = \rho(x, y)$. This translates directly into an invariance of the electric field upon this coordinate:

$$\mathbf{E}(x, y, z) = \mathbf{E}(x, y)$$

1.5.2 Invariance by spatial rotation around an axis

Consider the case of a charge distribution that is invariant under rotation along an axis (Oz for example). This means, in cylindrical coordinates,

$$\rho(r, \theta + a, z) = \rho(r, \theta, z) \quad \forall a \in \mathbb{R}$$

so that the charge density is independent of θ , $\rho = \rho(r, z)$, and this translates into an invariance of the electric field upon this coordinate:

$$\mathbf{E}(r, \theta, z) = \mathbf{E}(r, z)$$

Example 1.1 - Cylindrical/spherical symmetries

Two particular cases of relevant importance in electrostatics are those of a charge distribution having either cylindrical or spherical symmetry.

In the first case, the charge density ρ is invariant by spatial translation along Oz and also by spatial rotation around Oz . It is therefore independent of z and θ in cylindrical coordinates and one has

$$\rho(r, \theta, z) = \rho(r) \Rightarrow \mathbf{E}(r, \theta, z) = \mathbf{E}(r)$$

In the case of spherical symmetry, the charge density ρ is invariant under rotation around

any axis passing through the origin. It is then independent of θ and ϕ in spherical coordinates and this translates into:

$$\rho(r, \theta, \phi) = \rho(r) \Rightarrow \mathbf{E}(r, \theta, \phi) = \mathbf{E}(r)$$

Remark

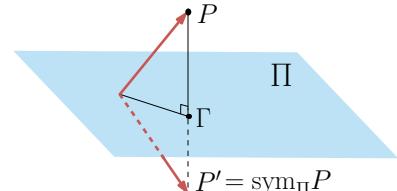
In conclusion, invariances of the charge distribution allows us to reduce the number of variables required to describe the spatial variation of \mathbf{E} . However, it does not tell us anything about the direction of the electric field. To constraint its direction, one can use the mirror symmetry or antisymmetry of the charge distribution as defined below.

1.5.3 Mirror symmetries

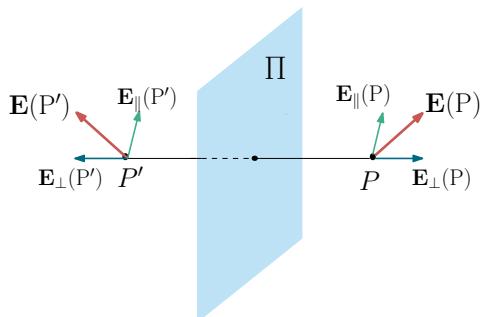
Let Π be a plane in space. For any point $P \in \mathbb{R}$ let us consider the reflection P' of P in the plane Π . In other words, P' is the symmetrical point of P in regard to the plane Π :

$$P' = \text{sym}_{\Pi} P$$

To find P' , first we determine the point $\Gamma \in \Pi$ which is closest to P . This determines a line passing through Γ and P , and $P' \neq P$ is the point on this line which is at the same distance to Γ than P .



- MIRROR SYMMETRY



We say that Π is a plane of symmetry for the charge distribution ρ if ρ satisfies the mirror symmetry with respect to Π :

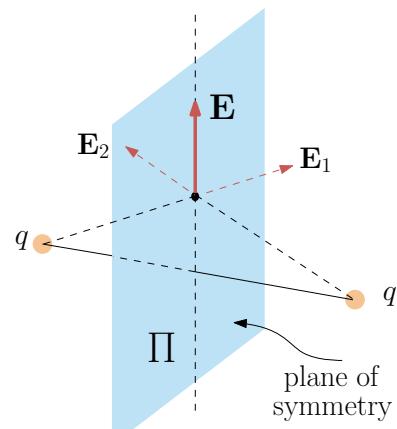
$$\rho(P') = \rho(P) \quad \text{with} \quad P' = \text{sym}_{\Pi} P$$

By Curie's principle, we conclude that the electric field generated by ρ exhibits this mirror symmetry as well

$$\mathbf{E}(P') = \text{sym}_{\Pi} \mathbf{E}(P)$$

The mirror symmetry for a vector quantity implies $\mathbf{E}_\perp(P) = -\mathbf{E}_\perp(P')$ and $\mathbf{E}_\parallel(P) = \mathbf{E}_\parallel(P')$. In particular, for a point $P \in \Pi$, one has $P' = P$ and therefore $\mathbf{E}_\perp(P) = -\mathbf{E}_\perp(P) = 0$. In other words:

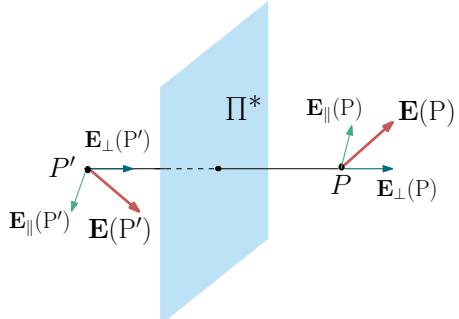
For any point P belonging to a symmetry plane Π , the electric field at P is within the plane Π .



Consider for example the plane Π equidistant to two identical charges q as shown in the figure below. This charge distribution has mirror symmetry with respect to Π . At any point in this plane, the superposition of the two electric fields generated separately by each charge gives indeed a total field within the plane, the normal components cancel each other out.

- MIRROR ANTISYMMETRY

We say that Π^* is a plane of antisymmetry for the charge distribution if ρ is antisymmetric with respect to Π^* :



$$\rho(P') = -\rho(P) \quad \text{with} \quad P' = \text{sym}_{\Pi} P$$

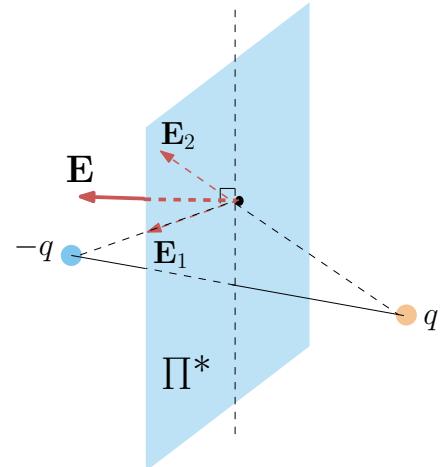
by Curie's principle, then the electric field generated by ρ exhibits this mirror antisymmetry as well

$$\mathbf{E}(P') = -\text{sym}_{\Pi^*} \mathbf{E}(P)$$

The mirror antisymmetry implies $\mathbf{E}_{\perp}(P) = \mathbf{E}_{\perp}(P')$ and $\mathbf{E}_{\parallel}(P) = -\mathbf{E}_{\parallel}(P')$. In particular, for a point $P \in \Pi^*$, one has $P' = P$ and therefore $\mathbf{E}_{\parallel}(P) = -\mathbf{E}_{\parallel}(P) = 0$. In other words:

For a point P inside an antisymmetry plane Π^* , the electric field at P is perpendicular to the plane Π^* .

Consider for example the plane Π^* equidistant to two charges of equal strength but of opposite sign (electric dipole) shown in the figure below. This charge distribution has mirror antisymmetry with respect to Π^* . At any point in this plane, the superposition of the two electric fields generated separately by each charge gives indeed a total field perpendicular to this plane, the parallel components cancel out.

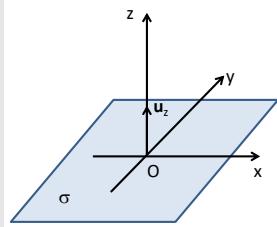


Remark

In conclusion, the orientation of the electric field at any point P is constrained by the symmetry (antisymmetry) planes of the charge distribution containing P . If such a plane exists, the electric field at P will be within (perpendicular) to this plane. The complete determination of the electric field and potential then consists in 1/finding its direction of the \mathbf{E} -field, 2/applying Gauss's theorem and 3/ finding the potential by integration of $\mathbf{E} = -\nabla V$.

1.6 Exercises

Exercise 1.1 - Field and potential for a uniformly charged plane



The infinite plane Oxy carries a uniform charge of surface density σ .

Calculate the electric field $\mathbf{E}(\mathbf{r})$ and the potential $V(\mathbf{r})$ at any point of space such that $z \neq 0$. For the potential, use the condition $V(0, 0, 0) = 0$.

Solution

The charge distribution is invariant by translation along the x - and the y -axes. thus the potential depends only on z and so does the electric field, which has a component along z only: $\mathbf{E}(\mathbf{r}) = -\nabla V(z) = -(\partial V / \partial z)\mathbf{u}_z$

$$\mathbf{E}(\mathbf{r}) = E(z)\mathbf{u}_z.$$

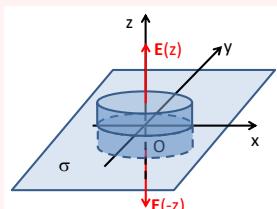
The direction of the field can also be obtained by considering that at any point $M(\mathbf{r})$, there are two symmetry planes for the charge distribution that contain M , one parallel to Oxz and the second parallel to Oyz . The electric field at M must be parallel to these two planes, hence, it must be in their intersection, that is along the z -axis.

The charge distribution is invariant by mirror symmetry with respect to the (Oxy) plane, thus, for two symmetric points M and M' , $\mathbf{E}(M') = -\mathbf{E}(M)$, that is $E(-z) = -E(z)$.

Apply Gauss's theorem considering as Gauss's surface a cylinder with basis of surface area S at $M(z)$ and $M'(-z)$, with $z > 0$. The electric field is uniform over the basis surfaces as it depends only on z . The flux of the electric field through the lateral surface is zero where since the field is parallel to the lateral surface, so the (outward) flux of \mathbf{E} through the cylinder is

$$E(z)S - E(-z)S = 2E(z)S = \frac{\sigma S}{\epsilon_0},$$

$E(-z) = -E(z)$ by symmetry, and the right-hand-side is the ratio of the total charge σS



enclosed in the Gauss surface, to ϵ_0 . The electric field finally reads

$$\mathbf{E}(z) = \frac{\sigma}{2\epsilon_0} \mathbf{u}_z, \quad \text{for } z > 0$$

and for $z < 0$, we find it by symmetry. The potential is found by integration using the condition that $V(0) = 0$,

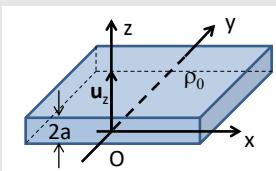
$$V(z) = -\text{sign}(z) \frac{\sigma}{2\epsilon_0} z.$$

The potential is non zero at infinity. This example with a simple geometry shows that when charges extend up to infinity, it may not be possible to fulfill a condition of zero potential at infinity. Here, neither \mathbf{E} nor V vanish at infinity. The alternative condition $V(0, 0, 0) = 0$ was chosen precisely for this reason.

Maxwell's equations tell us that fields and potentials (\mathbf{E} , V , \mathbf{B} , \mathbf{A}) do vanish at infinity if charges and currents are localized in a bounded region of space.

- In Maxwell's equations, \mathbf{E} , V , \mathbf{B} , \mathbf{A} assumed to vanish at infinity. If all charges ρ and currents \mathbf{j} are located at finite distance, then, V and \mathbf{A} vanish at infinity.
- For systems with simple geometry and charges at infinity \Rightarrow non zero potential at infinity
 \Rightarrow other convention.

Exercise 1.2 - Uniform volume charge between two planes



Consider the charge distribution defined by a uniform volume charge density between two horizontal planes:
If $z \in [-a, a]$, $\rho(M) = \rho_0$, otherwise, $\rho = 0$.
Express the electric field $\mathbf{E}(\mathbf{r})$ and the potential $V(\mathbf{r})$ at any point in space, assuming $V(0, 0, 0) = 0$.

Solution

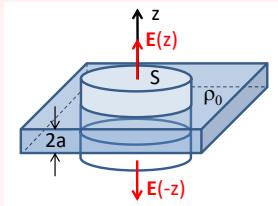
The charge distribution is invariant to translations along the x - or y -axis, thus the potential depends only on z : $V(\mathbf{r}) = V(z)$. From the relation $\mathbf{E} = -\nabla V$, we infer that the field is along the z axis:

$$\mathbf{E}(\mathbf{r}) = E(z) \mathbf{u}_z.$$

The plane $z = 0$ is a symmetry plane for the charge distribution, hence $E(-z) = -E(z)$ and it is sufficient to find an expression of the field and potential for $z > 0$ and then apply the symmetry.

We can apply Gauss's theorem to a cylinder of axis parallel to the z -axis and base surface S . There is no flux through the lateral surface as the field is parallel to it. The flux of the electric field is therefore equal to twice the flux through the top surface, that is $2E(z)S$ since the field is uniform on the top surface and has the same magnitude on the bottom surface.

Gauss's theorem then reads



$$2E(z)S = \frac{\rho_0}{\epsilon_0} \begin{cases} 2aS, & z \geq a \\ 2zS, & 0 \leq z \leq a \end{cases} \Rightarrow E(z) = \frac{\rho_0}{\epsilon_0} \begin{cases} a & z \geq a \\ z & 0 \leq z \leq a \end{cases}$$

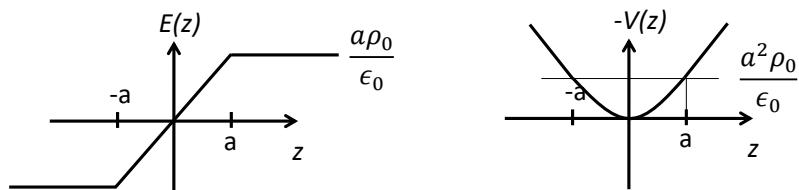
By integration ($V(z) = - \int^z E(z') dz'$), we find the potential

$$V(z) = -\frac{\rho_0}{\epsilon_0} \begin{cases} az + \text{Cst} & \text{if } z \geq a \\ z^2/2 + \text{Cst}' & \text{if } 0 \leq z \leq a \end{cases}$$

To find the integration constants, we express the continuity of the potential at $z = a$ and the condition $V(z = 0) = 0$:

$$\text{B.C.} \quad \begin{cases} V(z = 0) = 0 \Rightarrow \text{Cte}' = 0 \\ V(a) = -\frac{\rho_0}{\epsilon_0} a^2 + \text{Cte} = -\frac{\rho_0}{\epsilon_0} \frac{a^2}{2} \end{cases} \quad V(z) = -\frac{\rho_0}{\epsilon_0} \begin{cases} z^2/2 & \text{if } 0 \leq z \leq a \\ a^2/2 - az & \text{if } z > a \end{cases}$$

Finally, the field and potentials are plotted for $z > 0$ and the symmetry is applied to for the field and potentials for $z < 0$. There is a non-zero field and potential at infinity, which is due to the fact that the charge distribution is not bounded.



1.7 Summary and essential formulas

- Maxwell's equations form a complete set of coupled equations relating the electric and magnetic fields to their sources (charge ρ and current density \mathbf{J}):

$\nabla \cdot (\epsilon_0 \mathbf{E}) = \rho$	Gauss-Maxwell-
$\nabla \cdot \mathbf{B} = 0$	Thomson-Maxwell
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday-Maxwell
$\nabla \times \frac{\mathbf{B}}{\mu_0} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$	Ampère-Maxwell

which, together with the Lorentz force acting on a point charge q with velocity \mathbf{v}

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

provide a complete description of all classical electromagnetism.

- If the electromagnetic field satisfies Thomson-Maxwell's and Faraday-Maxwell's equations, a scalar potential $V(\mathbf{r}, t)$ and a vector potential $\mathbf{A}(\mathbf{r}, t)$ exist that are related to the electromagnetic field by

$$\begin{aligned}\mathbf{E}(\mathbf{r}, t) &= -\nabla V(\mathbf{r}, t) - \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} \\ \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t)\end{aligned}$$

- Since the electric charge is a conserved quantity, the charge density ρ and the current density \mathbf{J} are related at every point and at any time by the charge conservation law

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

- Gauss's theorem states that the flux of \mathbf{E} over any closed surface $\partial\Omega$ is proportional to the charge contained in the volume Ω enclosed by $\partial\Omega$

$$\iint_{\partial\Omega} \epsilon_0 \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{n}(\mathbf{r}) dS(\mathbf{r}) = Q_\Omega(t) = \iiint_\Omega \rho(\mathbf{r}, t) d^3 r.$$

- In electrostatics, the electric field is entirely determined by the two Maxwell's equations

$$\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}) &= \mathbf{0} & \Leftrightarrow & & \mathbf{E}(\mathbf{r}) &= -\nabla V(\mathbf{r}) \\ \nabla \cdot (\epsilon_0 \mathbf{E}(\mathbf{r})) &= \rho(\mathbf{r})\end{aligned}$$

- The electric potential V satisfies Poisson's equation $\nabla^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0}$ which, in regions free of charges, reduces to Laplace's equation $\nabla^2 V = 0$.
- Poisson's equation $\nabla^2 V(\mathbf{r}) = -\rho(\mathbf{r})/\epsilon_0$ in a finite volume Ω admits boundary conditions for which the solution is unique. These are:
 - Dirichlet's boundary condition: $V = V_D$ at any point of the boundary $\partial\Omega$.
 - Neumann's boundary condition: $E_\perp \equiv \mathbf{E} \cdot \mathbf{n} = E_N$ at any point in $\partial\Omega$.
- The electric field generated by a charge density $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ is given by the Coulomb integral

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{r}')(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} d^3 r'$$

This equation is valid when ρ represents the charge density of a continuous medium, and also in the case of a discrete distribution of charges.

The electric field derives from a scalar potential V such that $\mathbf{E} = -\nabla V$, which writes

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' + C$$

where C is a constant. If the charge density is integrable ($\iiint_{\mathbb{R}^3} |\rho(\mathbf{r})| d^3 r < \infty$), one has $C = \lim_{|\mathbf{r}| \rightarrow \infty} V(\mathbf{r})$.

- Thanks to the Dirac distribution, the charge density associated with a point charge q located at \mathbf{r}_0 can be written as a volume density $\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{r}_0)$. In this way, we obtain the electric field and the potential of a point charge (Coulomb's law):

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} q\delta(\mathbf{r}' - \mathbf{r}_0) \frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} d^3 r', \quad \text{i.e.,} \quad \mathbf{E}(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{(\mathbf{r} - \mathbf{r}_0)}{|\mathbf{r} - \mathbf{r}_0|^3}$$

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} q\delta(\mathbf{r}' - \mathbf{r}_0) \frac{1}{|\mathbf{r} - \mathbf{r}'|} d^3 r', \quad \text{i.e.,} \quad V(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{1}{|\mathbf{r} - \mathbf{r}_0|}$$

- Symmetry properties:

- Translation invariance eliminates dependency upon space coordinate.
- Rotation invariance around an axis eliminates dependency upon rotation angle.
- Symmetries: $\mathbf{r} \in (P)$, plane of: $\begin{cases} \text{Symmetry } \mathcal{S} \rightarrow \mathbf{E}(\mathbf{r}) \parallel (P). \\ \text{Antisymmetry } \mathcal{S}^* \rightarrow \mathbf{E}(\mathbf{r}) \perp (P). \end{cases}$

Chapter 2

Electrostatics of charge distributions

Introduction

The present chapter deals with the concepts of potential, field and energy of a charge distribution. For an isolated charge in vacuum, we already know that Coulomb's law gives us an expression for the electric field and potential generated by the charge at any point of space,

$$\mathbf{E}(\mathbf{r}) = \frac{q}{4\pi\epsilon_0 r^2} \mathbf{u}_r, \quad \text{and} \quad V(\mathbf{r}) = \frac{q}{4\pi\epsilon_0 r},$$

where \mathbf{u}_r denotes the radial unit vector of the spherical coordinate system centered at the position of the charge.

For a set of discrete charges q_i at positions \mathbf{r}_i , the superposition principle tells us that the field (potential) at any point of space is simply the sum of the fields (potentials) generated by each charge at that point.

$$\mathbf{E}(\mathbf{r}) = \sum_i \frac{q_i(\mathbf{r} - \mathbf{r}_i)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_i|^3}, \quad \text{and} \quad V(\mathbf{r}) = \sum_i \frac{q_i}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_i|}.$$

For a continuous charge distribution, we obtain the field and potential in the same way, simply replacing the sum over discrete charges by an integral.

$$\mathbf{E}(\mathbf{r}) = \iiint_{\Omega} \frac{\rho(\mathbf{r}')(\mathbf{r} - \mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|^3} d^3r', \quad \text{and} \quad V(\mathbf{r}) = \iiint_{\Omega} \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} d^3r',$$

where Ω denotes the volume enclosing the charge distribution and $\rho(\mathbf{r}')$ denotes its volume density at position \mathbf{r}' .

In this chapter, we will see what happens when the charge distribution is globally neutral, that is, $\sum_i q_i = 0$ or $\iiint_{\Omega} \rho(\mathbf{r}') d^3r' = 0$. We already know that a set of two opposite point charges, $\pm q$ at \mathbf{r}_{\pm} , separated by a distance a constitutes a globally neutral charge distribution called a dipole, characterized by its dipole moment $\mathbf{p} \equiv q\mathbf{a}$, where $\mathbf{a} \equiv \mathbf{r}_+ - \mathbf{r}_-$. In this chapter, we will generalize the concept of dipole to any neutral charge distribution.

The concept of dipole is extremely important in electromagnetism and more generally in physics as atoms and molecules are globally neutral objects which are constituents of the globally neutral matter surrounding us. Hence, from the point of view of the matter response to an electric field, neutral matter can be seen as an ensemble of dipoles which undergo displacements and reorientation, and which in turn generate electric fields everywhere in space, in particular acting on other dipoles. Interactions between dipoles, dipole layers, and more generally dipole distributions, lead to numerous applications in biophysics, chemical physics, condensed matter physics or plasma physics. For instance, remote sensing schemes based on dipole interactions are being developed to detect buried land mines and hidden explosives.

The potential and the field generated by a two-charge dipole at any point \mathbf{r} of space reads

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p} \cdot \mathbf{r}}{r^3}, \quad \text{and} \quad \mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r^5} (3(\mathbf{p} \cdot \mathbf{r})\mathbf{r} - r^2 \mathbf{p}),$$

where \mathbf{p} denotes the dipole moment located at the origin of the coordinate system. In comparison with the Coulomb potential that decays as $1/r$, the potential for a dipole decays faster, as $1/r^2$, while the field magnitude decays as $1/r^3$. In this chapter, we will generalize the concept of dipole and dipole moment to any neutral charge distribution. In this aim, we will look for the long range behavior of the potential (field) generated at large distances compared to the typical size of the charge distribution. This long range behavior is obtained by performing a multipole expansion of the potential (field).

• OUTLINE

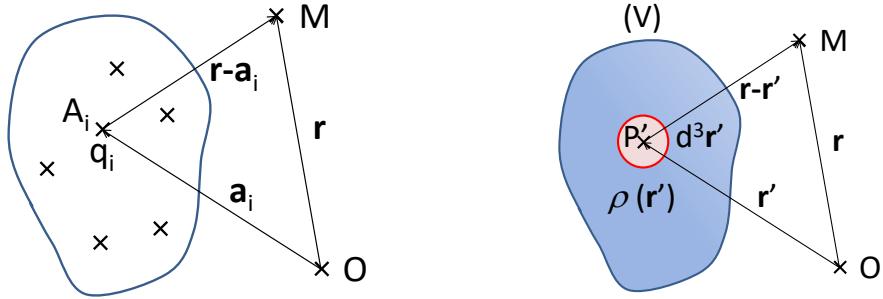
From the general expressions for the potential and field for any charge distribution, we will look for the potential and fields far from the charge distribution, leading to the different terms of the multipole expansion. We will then identify the different terms in the expansion as the monopole, dipole, quadrupole, ... contributions.

Interactions between dipoles requires the knowledge of the force and torque applied to a dipole charge distribution under the action of an electric field, which can itself be generated by a dipole distribution. A convenient way to access the force and torque applied to a dipole distribution is to derive these quantities from the energy of the charge distribution. Hence we will review the concept of energy for a charge distribution before deriving force and torques applied to a dipole distribution.

2.1 General expressions for the field and potential generated by a charge distribution

Let us suppose an arbitrary charge distribution inside a bounded region of space Ω . We can choose the origin of a coordinate system anywhere in space. Whether we consider a set of discrete charges at positions \mathbf{a}_i , $i = 1, 2, \dots$, or a continuous distribution of charges characterized

by its density $\rho(\mathbf{r}')$, we know that Coulomb's law and the superposition principle give us formal expressions for the potential generated by the charge distribution at any point \mathbf{r} of space, inside or outside of Ω .



These expressions read

$$V(\mathbf{r}) = \sum_{i=1}^{+\infty} \frac{q_i}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{a}_i|}$$

$$V(\mathbf{r}) = \iiint_{(V)} \frac{\rho(\mathbf{r}') d^3 \mathbf{r}'}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}$$

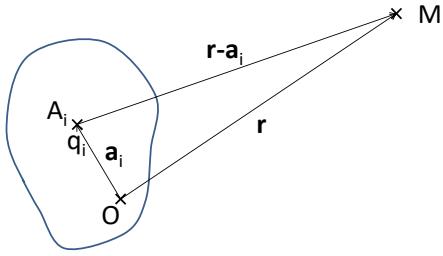
and can be seen as a single expression since a discrete charge distribution can be viewed as a continuous charge distribution of density $\rho(\mathbf{r}') = \sum_i q_i \delta(\mathbf{r}' - \mathbf{a}_i)$, where $\delta(\cdot)$ denotes the Dirac delta function. It is therefore sufficient to work with only one of these expressions, for instance in the case of a discrete charge distribution as the results can be generalized by transforming discrete sums into integrals involving a *continuous* density of charge distribution.

As for the electric field, Coulomb's law and the superposition principle allow us to find general expressions as reminded in introduction. However, we also recall that in electrostatics, the field and the potential are linked by the relation

$$\mathbf{E}(\mathbf{r}) = -\nabla V,$$

so that the knowledge of the potential $V(\mathbf{r})$ leads to the electric field at the price of three differentiations, which is often easier to perform compared to the volume integral arising from the combination of the superposition principle and Coulomb's field.

2.2 Multipole expansion of the potential for a discrete charge distribution



Let us perform the multipole expansion in the case of a discrete charge distribution. We assume that the size of the distribution is small with respect to the distance to the observation point. This condition is expressed by choosing an origin for the coordinate system that is close to the charge distribution. Then, all charge positions \mathbf{a}_i satisfy

$$|\mathbf{a}_i| \ll |\mathbf{r}| \quad \forall i,$$

that is, $\frac{a_i}{r}$ is infinitely small, which simply means that we look for the potential generated by the distribution of charges, far from it.

We will therefore introduce a Taylor expansion of $|\mathbf{r} - \mathbf{a}_i|^{-1}$ in the expression for the potential,

$$V(\mathbf{r}) = \sum_{i=1}^{+\infty} \frac{q_i}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{a}_i|}.$$

We start with a factorization of the leading order term ($1/r$) in the potential,

$$\frac{1}{|\mathbf{r} - \mathbf{a}_i|} = ((\mathbf{r} - \mathbf{a}_i)^2)^{-1/2} = (r^2 - 2\mathbf{a}_i \cdot \mathbf{r} + a_i^2)^{-1/2} = \frac{1}{r} \left(1 - 2\mathbf{a}_i \cdot \frac{\mathbf{r}}{r^2} + \frac{a_i^2}{r^2} \right)^{-1/2}.$$

We can now introduce the second order Taylor expansion, $(1+X)^\alpha \sim 1 + \alpha X + \alpha(\alpha-1)\frac{X^2}{2} + O(X^3)$, with $\alpha = -1/2$ and $X = -2\mathbf{a}_i \cdot \mathbf{r}/r^2 + a_i^2/r^2$, which yields, collecting orders up to the second,

$$\frac{1}{|\mathbf{r} - \mathbf{a}_i|} \sim \frac{1}{r} \left(1 + \frac{\mathbf{a}_i \cdot \mathbf{r}}{r^2} - \frac{a_i^2}{2r^2} + \frac{3(\mathbf{a}_i \cdot \mathbf{r})^2}{2r^4} \right).$$

Hence, separating different powers in a_i/r , we find the potential,

$$V(\mathbf{r}) \sim \frac{\sum_i q_i}{4\pi\epsilon_0 r} + \frac{\sum_i q_i \mathbf{a}_i \cdot \mathbf{r}}{4\pi\epsilon_0 r^3} + \frac{1}{2} \frac{\sum_i q_i (3(\mathbf{a}_i \cdot \mathbf{r})^2 - a_i^2 r^2)}{4\pi\epsilon_0 r^5}.$$

Each term of the multipole expansion varies as $1/r^n$, where $n = 1, 2, 3, \dots$, and as this is a Taylor expansion, higher orders $n > 3$ have been neglected. Physically, as we assumed that we look for the potential far from the distribution, the lowest order terms dominate in the expression for the potential. We can then identify the first three contributions.

• MONOPOLE (NON-NEUTRAL) DISTRIBUTION

The total charge of the distribution, also called the monopole moment, is non-zero,

$$Q \equiv \sum_i q_i \neq 0.$$

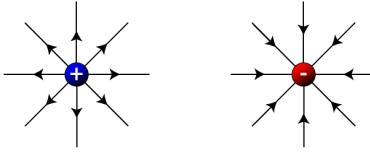
In that case, the potential is given by the first ($n = 1$) term in the multipole expansion. Neglecting higher-order terms, we find

$$V(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0 r}.$$

Far from the distribution, the potential is the same as the Coulomb potential generated by a single charge Q located at the origin O . Note that the result does neither depend on the location of the origin (provided it is close enough to the distribution) nor on the charge density or shape of the distribution.

The electric field found by applying $\mathbf{E} = -\nabla V$ also corresponds to the Coulomb field generated by a single charge Q equal to the total charge of the distribution:

$$\mathbf{E} = -\frac{\partial V}{\partial r} \mathbf{u}_r = \frac{Q}{4\pi\epsilon_0 r^2} \mathbf{u}_r$$



Field lines for a monopole positive (left) or negative (right) charge distribution.

The field lines look similar to those for a single charge and their orientation only depend on the sign of the total charge. Again, as we started from the result of an expansion, this result is valid only far from the distribution.

• DIPOLE DISTRIBUTION

In this case, $Q \equiv \sum_i q_i = 0$, so that the dominant lowest order term in the expansion of the potential corresponds to $n = 2$,

$$V(\mathbf{r}) \sim \sum_i \frac{q_i \mathbf{a}_i \cdot \mathbf{r}}{4\pi\epsilon_0 r^3},$$

which can be rewritten as

$$V(\mathbf{r}) = \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi\epsilon_0 r^3}, \quad \text{with} \quad \mathbf{p} = \sum_i q_i \mathbf{a}_i.$$

We recognize the same expression as that for the potential generated by a dipole made of two charges $\pm q$ at positions \mathbf{a}_{\pm} . The expression for the dipole moment \mathbf{p} now generalizes the already known expression $\mathbf{p} = q(\mathbf{a}_+ - \mathbf{a}_-)$ for the two-charge dipole. The expression for the potential

as a function of the dipole moment is the same whether we have a two-charge or a more general dipole.

The electric field for the dipole distribution is found by taking the gradient of the potential. This can be done either by differentiating the potential as a product of $\mathbf{p} \cdot \mathbf{r}$ by r^{-3} , or by expressing the potential in a basis of spherical coordinates and taking the gradient in this coordinate system.

In the first option, we note that if we have two functions $f(\mathbf{r})$ and $g(\mathbf{r})$, a formula of vectorial analysis allows us to calculate the gradient of the product

$$\nabla f g = f \nabla g + g \nabla f.$$

Applying this formula to $f = \mathbf{p} \cdot \mathbf{r}$ and $g = r^{-3}$, we find

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r^5} (3(\mathbf{p} \cdot \mathbf{r})\mathbf{r} - r^2 \mathbf{p}),$$

which, not astonishingly, corresponds to the known formula for the electric field generated by a two-charge-dipole.

In the second option, using spherical coordinates with the direction of the dipole moment serving as the reference axis to measure the polar angle θ , we have

$$\nabla V = \frac{\partial V}{\partial r} \mathbf{u}_r + \frac{1}{r} \frac{\partial V}{\partial \theta} \mathbf{u}_\theta + \frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi} \mathbf{u}_\phi.$$

Rewriting $\mathbf{p} \cdot \mathbf{r} = pr \cos \theta$, the potential depends only on r and θ ,

$$V(r, \theta) = \frac{p \cos \theta}{4\pi\epsilon_0 r^2}.$$

The electric field therefore reads

$$\mathbf{E}(\mathbf{r}) = E_r \mathbf{u}_r + E_\theta \mathbf{u}_\theta = -\frac{\partial V}{\partial r} \mathbf{u}_r - \frac{1}{r} \frac{\partial V}{\partial \theta} \mathbf{u}_\theta.$$

Finally,

$$\mathbf{E}(\mathbf{r}) = \frac{p}{4\pi\epsilon_0 r^3} (2 \cos \theta \mathbf{u}_r + \sin \theta \mathbf{u}_\theta).$$

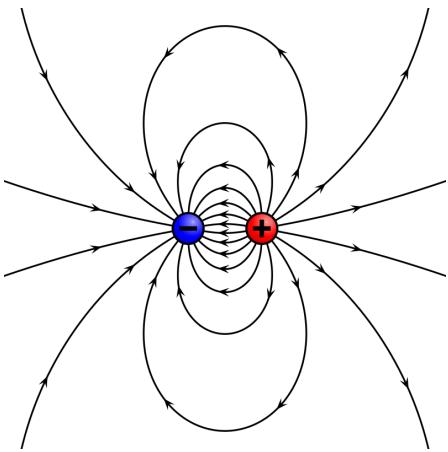
Far from the distribution, the field exhibits revolution symmetry around the axis of the dipole moment. Therefore, it is sufficient to show the field lines in a meridian plane $\phi = 0$ as the pattern is azimuthally symmetric, i.e., the same for all angle ϕ . The field lines are found from the differential equation (obtained by the condition $\mathbf{E} \times d\mathbf{r} = \mathbf{0}$, expressing that the electric field is tangent at any point \mathbf{r} to the field line, whose tangent vector is aligned with $d\mathbf{r}$)

$$\frac{rd\theta}{E_\theta} = \frac{dr}{E_r} \Rightarrow \frac{2d\theta}{\tan \theta} = \frac{dr}{r}.$$

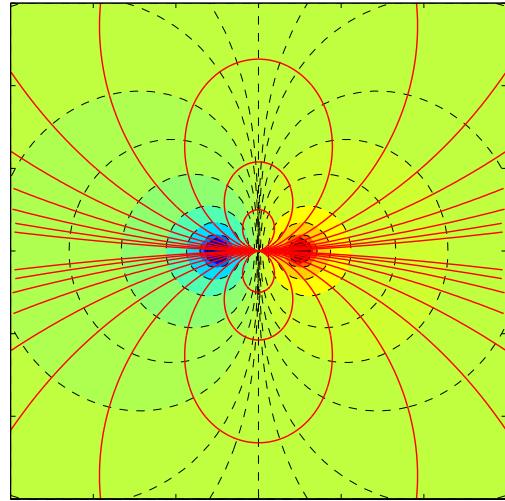
This equation has separated variables and is integrated as

$$\ln \sin^2 \theta = \ln r - \ln R_0 \quad \Rightarrow \quad r = R_0 \sin^2 \theta$$

where R_0 is an integration constant characterizing a field line. Different values for R_0 give different field lines (see figure).



Field lines for a dipole distribution represented as a two-charge-dipole.



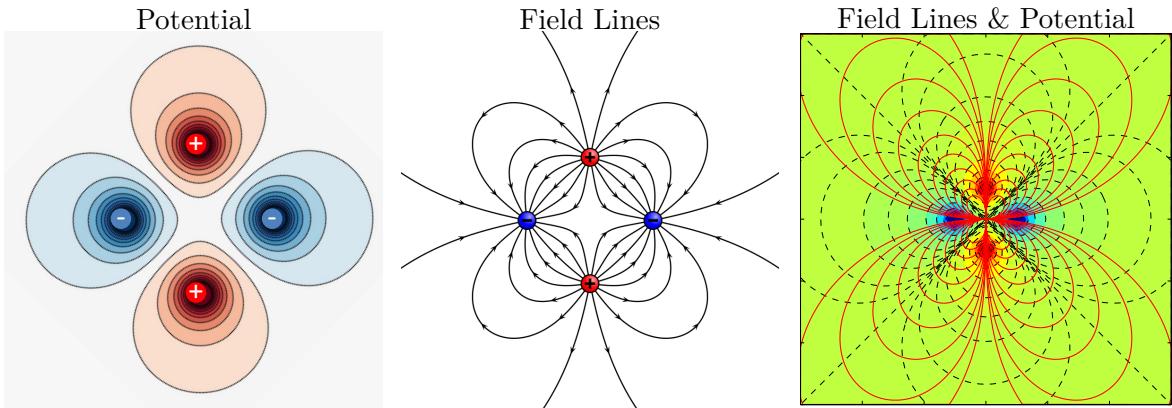
Field lines (red curves) and equipotentials (dashed lines) for a dipole distribution.

• QUADRUPOLE DISTRIBUTION

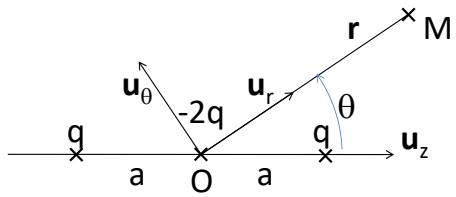
For the dipole distribution, we retained the lowest order $n = 2$ in the expansion for the potential, implicitly assuming that $\mathbf{p} = \sum_i q_i \mathbf{a}_i \neq 0$. If this is not the case, that is, if $\sum_i q_i \mathbf{a}_i = 0$, we must consider the next term in the expansion. If both the total charge and the dipole moment of the distributions are zero, $Q = \sum_i q_i = 0$ and $\mathbf{p} = \sum_i q_i \mathbf{a}_i = \mathbf{0}$, the lowest order non-zero term in the multipole expansion for the potential reads

$$V(\mathbf{r}) = \sum_i \frac{3q_i(\mathbf{a}_i \cdot \mathbf{r})^2 - q_i a_i^2 r^2}{8\pi\epsilon_0 r^5}.$$

It is equivalent to the potential for a distribution with three or four charges, called a quadrupolar distribution.



Example 2.1 - Quadrupolar distribution



Consider the set of three charges located on the z -axis, $+q$ at $z = \pm a$, and $-2q$ at $z = 0$. Using spherical coordinates, find the potential $V(r, \theta)$ at M , at 2nd order in $\frac{a}{r} \ll 1$. Find the electric field $\mathbf{E} = E_r \mathbf{u}_r + E_\theta \mathbf{u}_\theta$ and the field lines.

Solution

The total charge of the distribution is $Q = +q - 2q + q = 0$. The total dipole moment is

$$\mathbf{p} = \sum_i q_i \mathbf{a}_i = q(-a \mathbf{u}_z) - 2q(0 \mathbf{u}_z) + q(a \mathbf{u}_z) = \mathbf{0}$$

Thus, the charge distribution is quadrupolar and the potential can be found using the formula $V = \frac{1}{8\pi\epsilon_0 r^5} \sum_i q_i ((3\mathbf{a}_i \cdot \mathbf{r})^2 - a_i^2 r^2)$. Thus,

$$V = \frac{1}{8\pi\epsilon_0 r^5} [q(3(\mathbf{a} \cdot \mathbf{r})^2 - a^2 r^2) + q(3(-\mathbf{a} \cdot \mathbf{r})^2 - a^2 r^2) - 2q(3 \cdot 0 - \cdot 0)],$$

which yields

$$V(r, \theta) = \frac{1}{8\pi\epsilon_0 r^5} q(6(\mathbf{a} \cdot \mathbf{r})^2 - 2a^2 r^2) = \frac{qa^2}{4\pi\epsilon_0 r^3} (3 \cos^2 \theta - 1).$$

The electric field components are obtained from the relation

$$\mathbf{E} = -\nabla V = -\frac{\partial V}{\partial r} \mathbf{u}_r - \frac{1}{r} \frac{\partial V}{\partial \theta} \mathbf{u}_\theta.$$

We find

$$E_r = \frac{3qa^2}{4\pi\epsilon_0 r^4} (3 \cos^2 \theta - 1) \quad E_\theta = \frac{qa^2}{4\pi\epsilon_0 r^4} 6 \sin \theta \cos \theta.$$

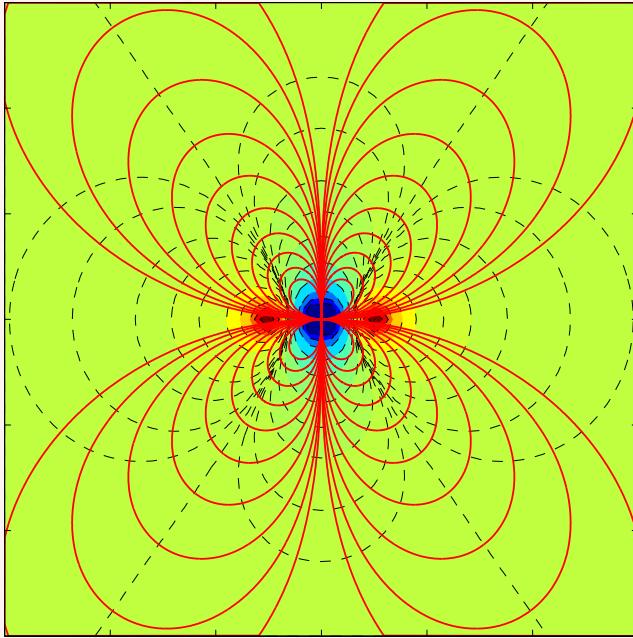
The field lines are found by solving the differential equation obtained from the relation $\frac{dr}{E_r} = \frac{rd\theta}{E_\theta}$, that is, after separation of variables,

$$\frac{dr}{r} = \frac{(3 \cos^2 \theta - 1)}{2 \sin \theta \cos \theta} d\theta = \left(\frac{\cos 2\theta}{\sin 2\theta} + \frac{\cos \theta}{2 \sin \theta} \right).$$

Integrating, we find

$$\ln \frac{r}{R_0} = \frac{1}{2} \ln |\sin 2\theta| + \frac{1}{2} \ln |\sin \theta| \Rightarrow r^2 = R_0^2 |\sin 2\theta \sin \theta|.$$

Equipotentials (dashed curves obtained for different values of \tilde{V}_0) and field lines (red curves obtained for different values of R_0) are shown in a meridian plane in the figure below.



Equipotential:

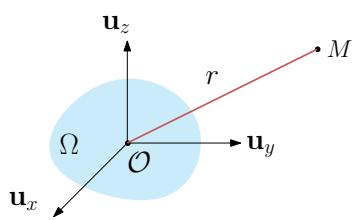
$$\tilde{V}_0 r^3 = (3 \cos^2 \theta - 1).$$

$$\text{with } \tilde{V}_0 \equiv V_0 \frac{4\pi\epsilon_0}{qa^2}.$$

Field lines:

$$r^2 = R_0^2 |\sin 2\theta \sin \theta|.$$

2.3 Multipole expansion of the potential for a continuous charge distribution



Let us now generalize the previous results to the case of an arbitrary charge distribution inside a bounded region of space Ω near the origin. Here, we will obtain an approximate expression for the

electric potential at any point M

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r',$$

far away from Ω . At distances $r = |\mathbf{r}|$ much larger than the typical size of Ω , we have $|\mathbf{r}'| \ll |\mathbf{r}|$. We first factorize $1/|\mathbf{r}| \equiv 1/r$,

$$|\mathbf{r} - \mathbf{r}'|^{-1} = \left(|\mathbf{r}|^2 - 2\mathbf{r} \cdot \mathbf{r}' + |\mathbf{r}'|^2 \right)^{-1/2} = \frac{1}{|\mathbf{r}|} \left(1 - 2 \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}|^2} + \frac{|\mathbf{r}'|^2}{|\mathbf{r}|^2} \right)^{-1/2}$$

and then, using variable $X \equiv -2 \frac{\mathbf{r} \cdot \mathbf{r}'}{r^2} + \frac{|\mathbf{r}'|^2}{r^2}$ and the Taylor expansion,

$$(1 + X)^{-1/2} = 1 - \frac{1}{2}X + \frac{3}{8}X^2 + O(X^3) \quad \text{for } X \ll 1,$$

we find,

$$\begin{aligned} \frac{1}{|\mathbf{r} - \mathbf{r}'|} &\approx \frac{1}{r} \left\{ 1 - \frac{1}{2} \left(\frac{|\mathbf{r}'|^2}{|\mathbf{r}|^2} - 2 \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}|^2} \right) + \frac{3}{8} \left(\frac{|\mathbf{r}'|^2}{|\mathbf{r}|^2} - 2 \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}|^2} \right)^2 \right\} \\ &= \frac{1}{r} \left\{ 1 + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^2} - \frac{1}{2} \frac{|\mathbf{r}'|^2}{r^2} + \frac{3}{2} \left(\frac{\mathbf{r} \cdot \mathbf{r}'}{r^2} \right)^2 + O\left(\frac{|\mathbf{r}'|^3}{r^3}\right) \right\}, \end{aligned}$$

so that the potential, at second order in $|\mathbf{r}'|/r$ is given by

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left\{ \frac{1}{r} \iiint_{\Omega} \rho(\mathbf{r}') d^3 r' + \frac{\mathbf{r}}{r^3} \cdot \iiint_{\Omega} \rho(\mathbf{r}') \mathbf{r}' d^3 r' + \frac{1}{2r^5} \iiint_{\Omega} (3(\mathbf{r} \cdot \mathbf{r}')^2 - |\mathbf{r}'|^2 r^2) \rho(\mathbf{r}') d^3 r' \right\}.$$

In the first term of the expansion, we recognize the total charge of the distribution, also called the electric monopole moment, whereas in the second we have the dipole moment. To factorize out the \mathbf{r} -dependent terms in the third term, the integrand can be rewritten

$$(3(\mathbf{r} \cdot \mathbf{r}')^2 - |\mathbf{r}'|^2 |\mathbf{r}|^2) = 3(\mathbf{r}^T \mathbf{r}') (\mathbf{r}'^T \mathbf{r}) - |\mathbf{r}'|^2 \mathbf{r}^T \mathbf{r} = \mathbf{r}^T (3\mathbf{r}' \mathbf{r}'^T - |\mathbf{r}'|^2) \mathbf{r},$$

where \mathbf{r}^T denotes the line vector obtained by transposition of the column vector \mathbf{r} . Hence, the product $\mathbf{r}' \mathbf{r}'^T$ is a matrix and the above expression involves only standard matrix vector products.

Definition 2.1: Dipole moments of an arbitrary charge distribution

For a continuous charge distribution of density ρ :

- The electric monopole moment Q writes

$$Q = \iiint_{\Omega} \rho(\mathbf{r}') d^3 r', \quad (2.1)$$

- The electric dipole moment \mathbf{p} writes

$$\mathbf{p} = \iiint_{\mathbb{R}^3} \mathbf{r} \rho(\mathbf{r}) d^3 r. \quad (2.2)$$

Note that for an electric dipole composed of two point charges $\pm q$ separated by a distance d , the dipole moment reduces to $\mathbf{p} = qd\mathbf{u}$, where \mathbf{u} is the unit direction joining the negative to the positive charge.

- The electric quadrupole moment is defined as the matrix

$$D = \iiint_{\Omega} \rho(\mathbf{r}') \left\{ 3\mathbf{r}' \cdot \mathbf{r}'^T - |\mathbf{r}'|^2 \right\} d^3 r'. \quad (2.3)$$

Finally, the multipole expansion writes

$$V(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0 r} + \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi\epsilon_0 r^3} + \frac{1}{8\pi\epsilon_0} \frac{\mathbf{r}^T D \mathbf{r}}{r^5} + \dots \quad (2.4)$$

which is a good approximation far from the charge distribution, that is, for $|\mathbf{r}| \gg \sup_{\mathbf{r}' \in \Omega} |\mathbf{r}'|$.

Remarks

1. The first term, that decays as $1/r$, corresponds to the potential of a point charge Q in the origin. Considering only this term in the expansion is the simplest approximation that one can make (if $Q \neq 0$).
2. The second term in the expansion, decaying as $1/r^2$, is approximately that of an electric dipole \mathbf{p} at the origin. Far from any neutral charge distribution ($Q = 0$), the potential is that of a dipole moment (if $\mathbf{p} \neq \mathbf{0}$).
3. The third term decays as $1/r^3$, and corresponds to the potential of a quadrupole. An elementary quadrupole is a set of 4 charges such that the total charge and total dipole moment are zero.

2.4 Expansion in Legendre polynomials

Let us suppose a region Ω in space free of charges ($\rho = 0$), such that the potential satisfies Laplace's equation

$$\nabla^2 V(\mathbf{x}) = 0, \quad \forall \mathbf{x} \in \Omega.$$

We will show that, if there is azimuthal symmetry, i.e., invariance of the potential with respect to a rotation around an axis, the general solution to Laplace's equation may be written as a series of well-known polynomial functions, called Legendre's polynomials. To show this, let us use spherical coordinates, and suppose that the potential is invariant to rotations around the z -axis. This means that the potential is independent of the ϕ coordinate, so that $V(r, \theta, \phi) = V(r, \theta)$. Now look for a solution of Laplace's equation of the form

$$V(r, \theta) = \frac{R(r)}{r} \Theta(\theta).$$

Injecting into Laplace's equation yields

$$\frac{\Theta}{r} \frac{d^2 R}{dr^2} + \frac{R}{r^3 \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = 0$$

and multiplying the latter equation by $\frac{r^3}{R(r)\Theta(\theta)}$, we obtain

$$\frac{r^2}{R} \frac{d^2 R}{dr^2} = -\frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right),$$

which is only possible if $\frac{r^2}{R} \frac{d^2 R}{dr^2}$ is equal to a constant¹, that we call λ . We then obtain an equation for Θ ,

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \lambda \Theta = 0,$$

whose solution is written in terms of Legendre polynomials P_l (see Appendix E for definition and properties of Legendre's polynomials). It can be shown that the regularity of Θ at $\theta = 0$ and $\theta = \pi$ imposes the condition $\lambda = l(l+1)$ where l is a positive integer, that is $l \in \mathbb{N}_0$. Then,

$$\Theta(\theta) = P_l(\cos \theta)$$

and the equation for the radial part of the potential writes

$$\frac{d^2 R}{dr^2} - \frac{l(l+1)}{r^2} R = 0,$$

whose solution is written:

$$R_l(r) = A_l r^{l+1} + B_l r^{-l}$$

with A_l and B_l constants.

¹as the left hand side of the equation to be solved is a function of r only, the right hand side is a function of θ only and the equation is valid for all r and θ

Definition 2.2: General solution of Laplace's equation with azimuthal symmetry

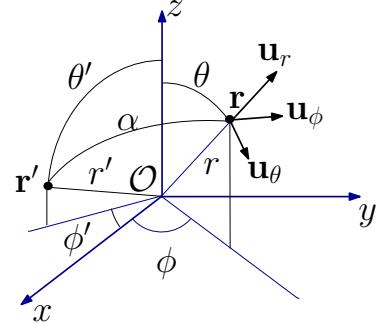
Let $\Omega \subseteq \mathbb{R}^3$ a region in space in which the potential satisfies Laplace's equation. If the potential is invariant to rotations around the z -axis, the potential can be written, in spherical coordinates, as a series involving the Legendre polynomials P_l ,

$$V(r, \theta) = \sum_{l=0}^{\infty} (A_l r^l + B_l r^{-(l+1)}) P_l(\cos \theta) \quad (2.5)$$

2.4.1 Multipole expansion with spherical coordinates

Rather than performing a Taylor expansion of $|\mathbf{r} - \mathbf{r}'|^{-1} = r^{-1}[1 - 2\mathbf{r} \cdot \mathbf{r}'(r'/r) + (r'/r)^2]^{-1/2}$ in the general expression for the electrostatic potential, it can be easier to look directly for an expansion of the potential written in terms of spherical coordinates. This expansion can be obtained by means of the identity

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} \sum_{l=0}^{+\infty} \left(\frac{r'}{r}\right)^l P_l(\mathbf{u}_r \cdot \mathbf{u}_{r'}) \quad \text{for } r' < r,$$



where $P_l(x)$ are the real-valued Legendre polynomials, which form an orthogonal and complete basis (see Appendix E), and the unit vectors are defined as $\mathbf{u}_r \equiv \frac{\mathbf{r}}{r}$ and $\mathbf{u}_{r'} \equiv \frac{\mathbf{r}'}{r'}$. The argument of $P_l(\mathbf{u}_r \cdot \mathbf{u}_{r'})$ is $\mathbf{u}_r \cdot \mathbf{u}_{r'} = \cos \alpha$ where α denotes the angle between \mathbf{r} and \mathbf{r}' . In the absence of any specific symmetry, it is convenient to express $P_l(\mathbf{u}_r \cdot \mathbf{u}_{r'})$ in terms of spherical harmonics $Y_{lm}(\theta, \phi)$, which form a complete and orthonormal basis (See Appendix F). This reads

$$P_l(\mathbf{u}_r \cdot \mathbf{u}_{r'}) = \frac{4\pi}{2l+1} \sum_{m=-l}^{m=+l} Y_{lm}(\theta', \phi') Y_{lm}(\theta, \phi).$$

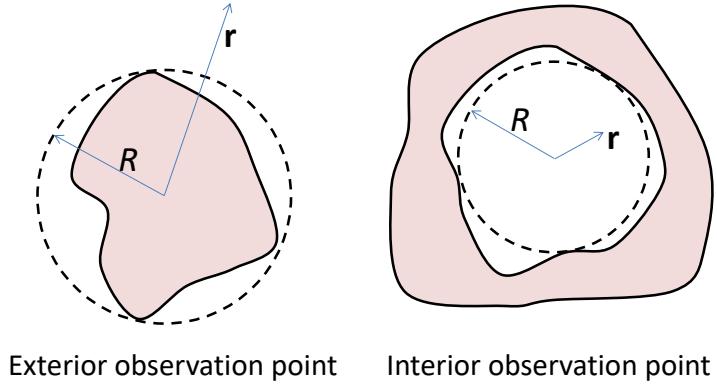
Substituting this expression into the inverse distance yields a spherical expansion for the inverse distance with a single origin of coordinates

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} \sum_{l=0}^{+\infty} \frac{4\pi}{2l+1} \left(\frac{r'}{r}\right)^l \sum_{m=-l}^l Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) \quad \text{for } r' < r.$$

Note that the role of primed and unprimed variables in the above expression can be exchanged, producing an expansion of the inverse distance that is valid when $r' > r$,

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r'} \sum_{l=0}^{+\infty} \frac{4\pi}{2l+1} \left(\frac{r}{r'}\right)^l \sum_{m=-l}^l Y_{lm}^*(\theta, \phi) Y_{lm}(\theta', \phi') \quad \text{for } r' > r.$$

The latter expansion is useful in situations where a charge distribution presents a cavity in which we are looking for an interior expansion of the potential.



For charge distributions confined within a sphere of radius R , the *exterior* spherical expansion of the potential reads

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r} \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathcal{M}_{lm}^{\text{ext}} \frac{Y_{lm}(\theta, \phi)}{r^l} \quad \text{for } r > R,$$

where \mathcal{M}_{lm} denote complex-valued spherical multipole moments of the charge distribution,

$$\mathcal{M}_{lm}^{\text{ext}} = \frac{4\pi}{2l+1} \iiint \rho(\mathbf{r}') r'^l Y_{lm}^*(\theta', \phi') d^3 r'.$$

For charge distributions surrounding a cavity, where now the cavity is confined within a sphere of radius R , the *interior* spherical expansion of the potential reads

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0 r} \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathcal{M}_{lm}^{\text{int}} \frac{Y_{lm}^*(\theta, \phi)}{r^l} \quad \text{for } r > R,$$

where the multipole moments read

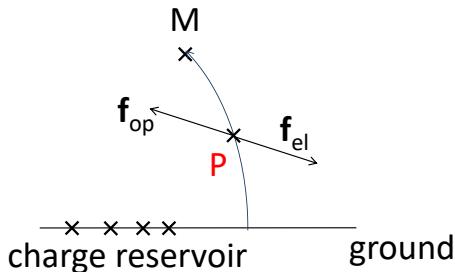
$$\mathcal{M}_{lm}^{\text{int}} = \frac{4\pi}{2l+1} \iiint \frac{\rho(\mathbf{r}')}{r'^{l+1}} Y_{lm}(\theta', \phi') d^3 r'.$$

For problems with azimuthal symmetry, $\rho(\mathbf{r}')$ does not depend on ϕ' ; it depends on r' and θ' , only: $\rho(r', \theta')$. The exterior and interior spherical expansions of the potential then reads

$$V(r, \theta) = \frac{1}{4\pi\epsilon_0} \begin{cases} \sum_{l=0}^{\infty} \frac{A_l}{r^{l+1}} P_l(\cos \theta), & A_l = \iiint r'^l \rho(r', \theta') P_l(\cos \theta') d^3 r', \quad r > R \\ \sum_{l=0}^{\infty} B_l r^l P_l(\cos \theta), & B_l = \iiint \frac{1}{r'^{l+1}} \rho(r', \theta') P_l(\cos \theta') d^3 r', \quad r < R \end{cases}$$

2.5 Energy for a discrete charge distribution

2.5.1 Energy of a single charge in a preexisting field



A charge (point \$P\$) is moved by an operator in a quasi-static way along the curved path from the reservoir (ground) to its final position at point \$M\$.

The operator works so as to exert a mechanical force that is exactly opposite to the electric force applied to the charge in the external field, at any time during its path from infinity to \$M\$. In doing so, the charge is moved in a quasistatic way without gaining kinetic energy. All the work done by the operator is stored by the charge in the form of potential energy.

The electric force applied to the charge at position \$P\$ is

$$\mathbf{f}_{el} = q\mathbf{E}(P) = -q \nabla V(P) = -\nabla W(P), \quad \text{with } W(P) = qV(P).$$

Thus, the operator exerts the force

$$\mathbf{f}_{op} = +q \nabla V(P)$$

and the work done by the operator is

$$\mathcal{W}_{op} = \int_{\infty}^M \mathbf{f}_{op} \cdot d\mathbf{P} = \int_{\infty}^M \underbrace{q \nabla V \cdot d\mathbf{P}}_{qdV} = qV(M) - qV(\infty), \quad \text{with } V(\infty) = 0.$$

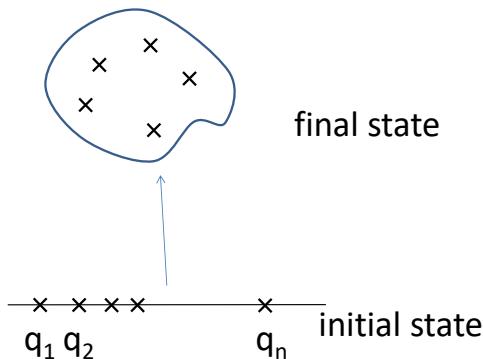
We find,

$$\mathcal{W}_{op} = qV(M).$$

The work of the operator does not depend on the path used to bring the charge at \$M\$.

The potential energy of a charge \$q\$ at potential \$V\$ is \$U_p = qV\$

2.5.2 Constitution energy for a discrete charge distribution



The charges are grounded in their initial state. An operator moves them one after another to their final position in the distribution.

The work to bring the third charge q_3 is $W_3 = q_3 V_{3(1,2)}$, where $V_{3(1,2)}$ denotes the potential at the position of q_3 generated by q_1 and q_2 . Proceeding with the entire set of charges one after another, the total work done by the operator, which is equal to the constitution (potential) energy U_p , reads

$$U_p = W = \sum_{i=1}^n W_i = 0 + q_2 V_{2(1)} + q_3 V_{3(2,1)} + \dots + q_n V_{n(n-1,\dots,1)},$$

where $V_{n(n-1,\dots,1)}$ denotes the potential at the position of q_n generated by all other charges.

Now let us calculate again the work of the operator to bring the charges in their final position, starting with charge q_n and proceeding in reverse order down to q_1 . The work to bring q_n in its final position is $W'_n = 0$ as there is no preexisting field. The work to bring q_{n-1} is $W'_{n-1} = q_{n-1} V_{n-1(n)}$. Proceeding down to q_1 , we find a second expression for the constitution energy U_p that must take the same value

$$U_p = W' = \sum_{i=1}^n W'_i = q_1 V_{1(n,\dots,2)} + q_2 V_{2(n,\dots,3)} + \dots + q_{n-1} V_{n-1(n)},$$

since it corresponds to the same charge distribution. Adding now the two expressions for U_p , we find

$$2U_p = q_1 V_1 + q_2 V_2 + \dots + q_{n-1} V_{n-1} + q_n V_n,$$

where

$$q_i V_i = W_i + W'_i, \quad \text{that is} \quad V_i = V_{i(i-1,\dots,1)} + V_{i(n,\dots,i+1)}.$$

In other words, V_i denotes the potential generated at the final position of charge q_i by all other charges held in their final positions. The result writes more compactly as

$$U_p = \sum_i \frac{q_i V_i}{2}.$$

2.5.3 Constitution energy for a continuous charge distribution

It is straightforward to generalize the formula obtained for a discrete charge distribution in the case of a continuous charge density $\rho(\mathbf{r})$ enclosed in a volume Ω ,

$$U_p = \iiint_{\Omega} \frac{\rho(\mathbf{r}') V(\mathbf{r}')}{2} d^3 r' = \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{r}') V(\mathbf{r}')}{2} d^3 r'. \quad (2.6)$$

Since the charge density is zero out of the volume Ω , it is possible to extend the integration volume to \mathbb{R}^3 .

2.5.4 Electrostatic energy density

From (2.6), we see that the electrostatic energy of a charge distribution ρ can be obtained as an integral of the form

$$U = \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{r}) V(\mathbf{r})}{2} d^3 r$$

The differential form of Gauss's law allows us to rewrite the charge density as $\rho = \epsilon_0 \nabla \cdot \mathbf{E}$, and so,

$$U = \iiint_{\mathbb{R}^3} \frac{\epsilon_0}{2} V(\mathbf{r}) (\nabla \cdot \mathbf{E}(\mathbf{r})) d^3 r$$

In addition, the divergence of the product of a scalar by a vector reads

$$\nabla \cdot (V \mathbf{E}) = V(\nabla \cdot \mathbf{E}) + \mathbf{E} \cdot \nabla V,$$

thus,

$$\begin{aligned} V(\nabla \cdot \mathbf{E}) &= \nabla \cdot (V \mathbf{E}) - \mathbf{E} \cdot (\nabla V) \\ &= \nabla \cdot (V \mathbf{E}) - \mathbf{E} \cdot (-\mathbf{E}) \\ &= \nabla \cdot (V \mathbf{E}) + |\mathbf{E}|^2. \end{aligned}$$

Finally, the electrostatic energy can be written as

$$U = \frac{\epsilon_0}{2} \iiint_{\mathbb{R}^3} (\nabla \cdot (V \mathbf{E}) + |\mathbf{E}|^2) d^3 r.$$

The first term corresponds to a surface integral (Green-Ostrogradsky's theorem) at infinity. It is null since V and \mathbf{E} both decay to zero at infinity if ρ is integrable. Finally,

$$U = \frac{\epsilon_0}{2} \iiint_{\mathbb{R}^3} |\mathbf{E}(\mathbf{r})|^2 d^3 r. \quad (2.7)$$

The electrostatic energy density (energy per unit volume) can be written as

$$u_E(\mathbf{r}) = \frac{\epsilon_0}{2} |\mathbf{E}(\mathbf{r})|^2.$$

The electrostatic energy can be therefore thought to be localized in those regions of space where there is an electric field. The total electrostatic energy U of a system is the integral

$$U = \iiint_{\mathbb{R}^3} u_E(\mathbf{r}) d^3 r.$$

2.6 Effect of an external electric field on a dipole distribution

Consider a dipole distribution of discrete charges q_i , placed at fixed positions \mathbf{r}_i in a static (time independent) electric field $\mathbf{E}(\mathbf{r})$. As a dipole distribution, it is neutral and satisfies $\sum_i q_i = 0$. Let us assume that the charge distribution cannot be deformed, that is, it is similar to a solid body undergoing forces $\mathbf{f}_i = q_i \mathbf{E}(\mathbf{r}_i)$ applied on each charge q_i .

The total force applied to the distribution is

$$\mathbf{F} = \sum_i \mathbf{f}_i.$$

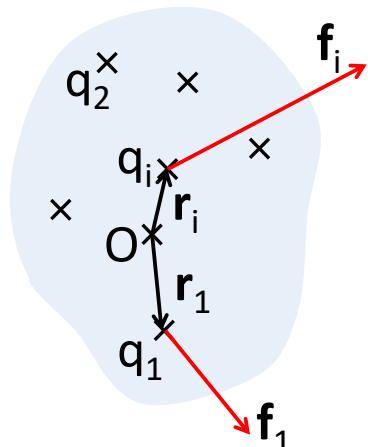
If we choose an origin O that coincides with the *position* of the charge distribution (for instance, O can be chosen as the center of mass), the torque at O then reads

$$\boldsymbol{\tau}(O) = \sum_i \mathbf{r}_i \times \mathbf{f}_i.$$

Our goal is to find expressions for \mathbf{F} and $\boldsymbol{\tau}(O)$ as functions of the dipole moment $\mathbf{p} = \sum_i q_i \mathbf{r}_i$.

- FORCE AND TORQUE USING THE INTERACTION ENERGY

Recall that for a charge q in an external potential V , the potential energy of the charge is given by the function $U_p = qV$. For a distribution of charges q_i at \mathbf{r}_i , the potential energy is



$U_p = \sum_i q_i V(\mathbf{r}_i)$. An expression of U_p as a function of the dipole moment is found by Taylor expanding the potentials $V(\mathbf{r}_i)$ around the position of the dipole $\mathbf{r}_O = \mathbf{0}$, chosen as the origin,

$$V(\mathbf{r}_i) = V(\mathbf{r}_O = \mathbf{0}) + [(\mathbf{r}_i \cdot \nabla) V(\mathbf{r})]|_{\mathbf{r}_O=\mathbf{0}}.$$

Hence, the potential energy of the dipole distribution reads

$$\begin{aligned} U_p(\mathbf{r}_O) &= \underbrace{\sum_i q_i V(\mathbf{r}_O)}_0 + \sum_i q_i [(\mathbf{r}_i \cdot \nabla) V(\mathbf{r})]|_{\mathbf{r}_O} \\ &= 0 + \underbrace{[(\sum_i q_i \mathbf{r}_i) \cdot \nabla V(\mathbf{r})]|_{\mathbf{r}_O}}_{\mathbf{p}} = -\mathbf{p} \cdot \mathbf{E}(\mathbf{r}_O) \end{aligned}$$

The potential (interaction) energy function for the charge distribution in the scalar potential $V(\mathbf{r})$ is therefore

$$U_p(\mathbf{r}) = -\mathbf{p} \cdot \mathbf{E}(\mathbf{r})$$

The energy of interaction $U_p = -\mathbf{p} \cdot \mathbf{E}(\mathbf{r})$ is a function of the position and orientation of the dipole (six degrees of freedom). It is minimum when the dipole moment \mathbf{p} aligns with the electric field \mathbf{E} . In order to derive the force \mathbf{F} and torque $\mathcal{T}(O)$ acting on the dipole, consider an elementary displacement $d\mathbf{r}$ and an elementary rotation $d\boldsymbol{\omega}$ from a state of mechanical equilibrium of the dipole. The related change in interaction energy reads

$$\begin{aligned} dU_p &= -\mathbf{p} \cdot d\mathbf{E} - d\mathbf{p} \cdot \mathbf{E}(\mathbf{r}) \\ &= -\mathbf{p} \cdot \underbrace{\left(\frac{\partial \mathbf{E}}{\partial x} dx + \frac{\partial \mathbf{E}}{\partial y} dy + \frac{\partial \mathbf{E}}{\partial z} dz \right)}_{d\mathbf{E}} - \underbrace{(d\boldsymbol{\omega} \times \mathbf{p}) \cdot \mathbf{E}}_{d\mathbf{p}}. \end{aligned}$$

This must be identified with the elementary work $dW = dU_p$ performed by an external operator acting on the dipole with force $-\mathbf{F}$ and torque $-\mathcal{T}(O)$ to maintain mechanical equilibrium, that is to say

$$dW = -\mathbf{F} \cdot d\mathbf{x} - \mathcal{T}(O) \cdot d\boldsymbol{\omega}.$$

Finally, identifying dW and dU_p , the components of the force acting on the dipole are

$$F_x = \mathbf{p} \cdot \frac{\partial \mathbf{E}}{\partial x}, \quad F_y = \mathbf{p} \cdot \frac{\partial \mathbf{E}}{\partial y}, \quad F_z = \mathbf{p} \cdot \frac{\partial \mathbf{E}}{\partial z},$$

and the torque is obtained by applying a circular shift to the scalar triple product $(d\boldsymbol{\Omega} \times \mathbf{p}) \cdot \mathbf{E} = (\mathbf{p} \times \mathbf{E}) \cdot d\boldsymbol{\Omega}$,

$$\mathcal{T}(O) = \mathbf{p} \times \mathbf{E}.$$

• DIRECT CALCULATION

The force acting on the dipole could have been obtained from a direct Taylor expansion of the fields $\mathbf{E}(\mathbf{r}_i)$ in the total force $\mathbf{F} = \sum_i q_i \mathbf{E}(\mathbf{r}_i)$. Consider for instance the x -component

$$\begin{aligned} F_x(\mathbf{r}_O) = \sum_i q_i E_x(\mathbf{r}_i) &= \sum_i q_i \left\{ E_x(\mathbf{r}_O = \mathbf{0}) + [(\mathbf{r}_i \cdot \nabla) E_x(\mathbf{r})] \Big|_{\mathbf{r}_O} \right\} \\ &= \underbrace{\left(\sum_i q_i \right)}_0 E_x(\mathbf{r}_O) + \underbrace{\left(\sum_i q_i \mathbf{r}_i \right)}_{\mathbf{p}} \cdot \nabla E_x \Big|_{\mathbf{r}_O} \\ &= (\mathbf{p} \cdot \nabla) E_x \end{aligned}$$

A similar calculation can be done for the other components, which yields $F_y(\mathbf{r}_O) = (\mathbf{p} \cdot \nabla) E_y(\mathbf{r}_O)$, $F_z(\mathbf{r}_O) = (\mathbf{p} \cdot \nabla) E_z(\mathbf{r}_O)$ and finally, a compact expression reads

$$\boxed{\mathbf{F}(\mathbf{r}) = (\mathbf{p} \cdot \nabla) \mathbf{E}(\mathbf{r})}. \quad (2.8)$$

As for the torque,

$$\mathcal{T}(O) = \sum_i \mathbf{r}_i \times \mathbf{f}_i = \sum_i \mathbf{r}_i \times q_i \mathbf{E}(\mathbf{r}_i),$$

a Taylor expansion of $\mathbf{E}(\mathbf{r}_i)$ around O writes $\mathbf{E}(\mathbf{r}_i) = \mathbf{E}(\mathbf{r}_O) + \dots$. Hence,

$$\mathcal{T}(O) = \left(\sum_i q_i \mathbf{r}_i \right) \times \mathbf{E}(\mathbf{r}_O) \Rightarrow \mathcal{T}(O) = \mathbf{p} \times \mathbf{E}(\mathbf{r}_O)$$

We retrieved the result

$$\boxed{\mathcal{T} = \mathbf{p} \times \mathbf{E}}$$

which coincides with the expression found by using the electrostatic interaction energy. At this stage, careful readers have noticed that we obtained two different expressions for the force acting on the dipole. Indeed, consider for instance the x -component of the force, the two expressions are equal only if

$$(\mathbf{p} \cdot \nabla) E_x(\mathbf{r}) = \mathbf{p} \cdot \frac{\partial \mathbf{E}(\mathbf{r})}{\partial x}.$$

Expanding the above equation, we find

$$p_x \frac{\partial E_x}{\partial x} + p_y \frac{\partial E_x}{\partial y} + p_z \frac{\partial E_x}{\partial z} = p_x \frac{\partial E_x}{\partial x} + p_y \frac{\partial E_y}{\partial x} + p_z \frac{\partial E_z}{\partial x},$$

or simply

$$p_y \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right) + p_z \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right) = 0,$$

which is true when $\nabla \times \mathbf{E} = \mathbf{0}$, i.e., for electrostatics, but not true in general. The different expressions therefore originate in the different assumptions made in the derivations: Deriving

the force acting on the dipole from the interaction energy of the dipole is valid only in the context of electrostatics, while the direct calculation did not require any approximation apart from the Taylor expansion performed in both cases.

The most general expression for the force acting on the dipole is finally(2.8).

This implies that a dipole will be pushed towards regions with larger electric fields. In the case where \mathbf{p} does not depend on the electric field, we can also write

$$\mathbf{F} = -\nabla U_p = \nabla (\mathbf{p} \cdot \mathbf{E}),$$

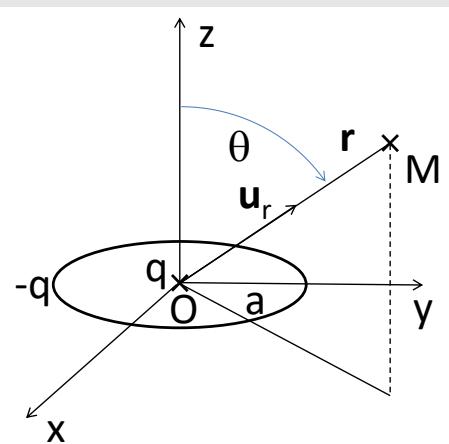
that is to say, the force is conservative. Finally, note that the torque

$$\mathcal{T} = \mathbf{p} \times \mathbf{E} \quad (2.9)$$

acts on a dipole even in the absence of a net force there. Indeed, even if the electric field is uniform (zero gradient, implies zero force), it is not necessarily aligned with the dipole moment, producing a non zero torque with respect to the center of gravity of the dipole.

2.7 Exercises

Exercise 2.1 - Uniform negative charge on a ring and positive point charge at the center



The charge distribution consists in a point charge $+q$ at the origin O , and a negative charge $-q$ uniformly distributed on the circle of center O and radius a in the plane xy .

- a.** What type of distribution is it?
- b.** Discuss the symmetries of the charge distribution and infer the dependence of the potential upon coordinates.
- c.** Calculate the potential $V(r, \theta)$ at $M(\mathbf{r})$ for $r \gg a$.
- d.** Plot equipotentials and field lines.

Solution

- a.** The total charge of the system is zero. If $\lambda = -q/(2\pi a)$ denotes the line charge density on the ring, the dipole moment reads

$$\mathbf{p} = q\mathbf{0} + \int_{\phi=0}^{\phi=2\pi} \lambda a \mathbf{u}_\rho,$$

where $\mathbf{u}_\rho = \cos \phi \mathbf{u}_x + \sin \phi \mathbf{u}_y$. Therefore, we find $\mathbf{p} = \mathbf{0}$. The charge distribution is a quadrupole.

- b.** The charge distribution is invariant to a rotation around the z -axis. Therefore, using spherical coordinates, the potential does not depend on the polar angle ϕ . We write $V(\mathbf{r}) = V(r, \theta)$. It is sufficient to find the potential in a meridian plane.
- c.** Far from a quadrupole distribution, the expression for the potential in the case of discrete charges of positions \mathbf{a}_i read

$$V = \sum_i q_i \frac{3(\mathbf{a}_i \cdot \mathbf{r})^2 - a_i^2 r^2}{8\pi\epsilon_0 r^5}.$$

The charge at $\mathbf{a}_i = \mathbf{0}$ does not contribute. The other charges consist in a superposition of charge elements $\lambda ad\phi$ at position $\mathbf{a}(\phi) = a\mathbf{u}_\rho(\phi)$, with $\lambda 2\pi a = -q$. Hence, the potential reads

$$\begin{aligned} V &= \int \lambda \frac{3(\mathbf{a} \cdot \mathbf{r})^2 - a^2 r^2}{8\pi\epsilon_0 r^5} ad\phi = \frac{\lambda a^2 r^2}{8\pi\epsilon_0 r^5} \int_{\phi=0}^{\phi=2\pi} (3 \sin^2 \theta \cos^2 \phi - 1) ad\phi \\ &= \frac{\lambda a^3}{8\pi\epsilon_0 r^3} \pi (3 \sin^2 \theta - 2) = \frac{-qa^2}{8\pi\epsilon_0 r^3} \left(\frac{3}{2} \sin^2 \theta - 1 \right). \end{aligned}$$

- d.** Equipotentials satisfy $V(r, \theta) = \text{Cst}$, that is

$$r(\theta) = C \left(\frac{3}{2} \sin^2 \theta - 1 \right)^{1/3}.$$

They are shown in blue curves (for different values of C) in the figure below.

The electric field is found from the relation $\mathbf{E} = -\nabla V$ and reads

$$\mathbf{E} = \frac{qa^2}{16\pi\epsilon_0 r^4} \left[-3(3 \sin^2 \theta - 2) \mathbf{u}_r + 6 \sin \theta \cos \theta \mathbf{u}_\theta \right].$$

Field lines are found by integrating the differential equation

$$\frac{dr}{E_r} = \frac{rd\theta}{E_\theta} \Rightarrow \frac{dr}{r} = \frac{2 - 3 \sin^2 \theta}{2 \sin \theta \cos \theta} = \frac{1}{2 \tan \theta} + \frac{1}{\tan 2\theta},$$

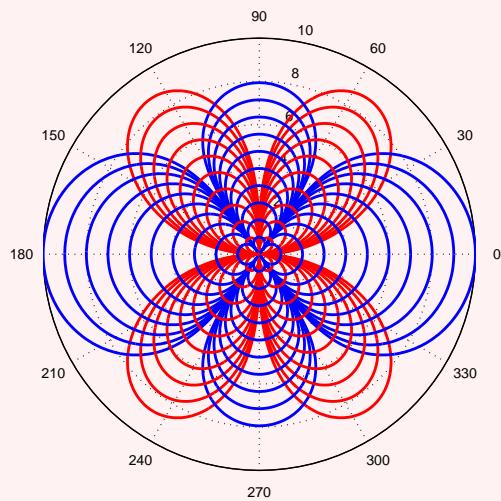
where we separated variables r and θ . Integrating, we find

$$\ln \frac{r}{R_0} = \frac{1}{2} \ln |\sin \theta| + \frac{1}{2} \ln |\sin 2\theta|$$

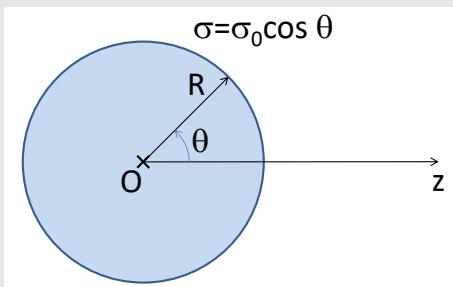
Finally the field lines satisfy the equation

$$r^2 = R_0^2 |\sin 2\theta \sin \theta|$$

and are plotted as red curves in the figure below. We recognize the pattern for a quadrupole.



Exercise 2.2 - Field at the center of a charged sphere



Calculate the electric field at the center of a sphere carrying a surface charge of density $\sigma = \sigma_0 \cos \theta$.

Solution

Any plane containing the z -axis is a symmetry plane of the charge distribution, hence the electric field at O must be in all of these planes, so it is along the z -axis. Let's use the superposition principle and consider the charge distribution as a superposition of spherical rings centered on the z -axis and described by polar angles between θ and $\theta + d\theta$. The corresponding element of surface area is $dS = 2\pi R^2 \sin \theta d\theta$ and the element of charge on this surface, at distance R from O , is $dq = \sigma dS$.

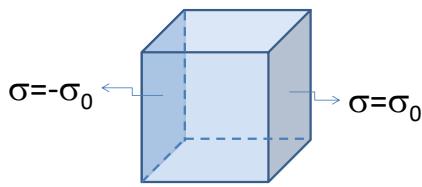
Each surface element of the ring contributes to a field element at O directed along $-\mathbf{u}_r$, so in projection along the z -axis, the charge element of the ring dq contributes to the z -component of the field at O with the field element

$$dE_z = \frac{dq}{4\pi\epsilon_0 R^2} (-\mathbf{u}_r \cdot \mathbf{u}_z) = -\frac{\sigma_0 \cos^2 \theta \sin \theta d\theta}{2\epsilon_0}.$$

Integrating over θ , we find the field at O ,

$$\mathbf{E} = -\frac{\sigma_0}{2\epsilon_0} \int_0^\pi \cos^2 \theta \sin \theta d\theta = -\frac{\sigma_0}{2\epsilon_0} \left[\frac{\cos^3 \theta}{3} \right]_0^\pi = \frac{\sigma_0}{3\epsilon_0} \mathbf{u}_z.$$

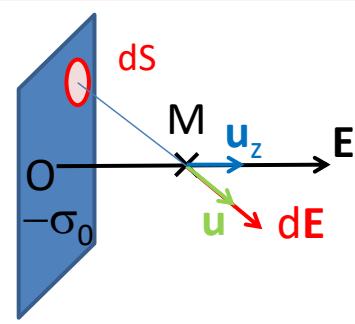
Exercise 2.3 - Field at the center of a cube



The left face of the cube carries a surface charge of density $-\sigma_0$. The opposite (right) face carries a surface charge of density $+\sigma_0$. All other faces carry no charge.

Calculate the field at the center M of the cube.

Solution



The field at M is along the z -axis, the symmetry axis of the cube that connect the centers of the two charged faces.

Consider first a single face, for instance the left face. A surface element dS on this face generates an electric field at M whose projection along z reads $dE_z = d\mathbf{E} \cdot \mathbf{u}_z$, with

$$dE_z = \frac{\sigma_0 dS}{4\pi\epsilon_0 r^2} \mathbf{u} \cdot \mathbf{u}_z$$

where r is the distance of the charge element to M and \mathbf{u} is the unit vector pointing from the charge element to M . In this expression, we remark that $\frac{dS}{r^2} \mathbf{u} \cdot \mathbf{u}_z$ is exactly the element of solid angle $d\Omega$ subtended by dS at M .

$$dE_z = \frac{\sigma_0}{4\pi\epsilon_0} d\Omega$$

If M is at the center of the cube, the solid angle subtended by the left face is $\Omega = \frac{4\pi}{6}$, that is, $1/6$ of the total solid angle of 4π for the entire cube.

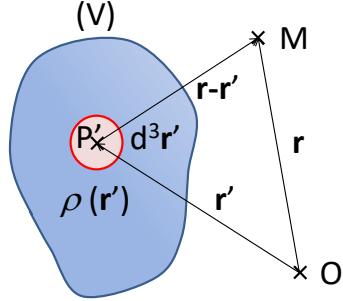
Now the right face will contribute to the z -component of the electric field in the same direction as for the left face because the opposite charge is associated with an opposite direction from the charge to M . Thus, for the two faces, we find twice the contribution of the left face.

$$\mathbf{E} = 2 \times \frac{\sigma_0}{4\pi\epsilon_0} \Omega \mathbf{u}_z = \frac{\sigma_0}{3\epsilon_0} \mathbf{u}_z.$$

The result is the same for the cube as for the sphere. We will use this result to investigate the concept of local field in chapter 5.

2.8 Summary and essential formulas

For a charge distribution of volume density $\rho(\mathbf{r})$, the electrostatic potential reads



$$V(\mathbf{r}) = \iiint_{(V)} \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|} d^3 r', \quad \mathbf{r}' = \mathbf{OP}'$$

The moments of the charge distribution are defined as:

$$\text{Monopole moment: } Q = \iiint_{(V)} \rho(\mathbf{r}') d^3 r'$$

$$\text{Dipole moment: } \mathbf{p} = \iiint_{(V)} \rho(\mathbf{r}') \mathbf{r}' d^3 r'$$

- Far from a charge distribution, that is for $r \gg a$ where a denotes the largest dimension of the distribution, we can perform a multipole expansion of the potential, which is an expansion in powers of a/r .

$$V(\mathbf{r}) = \underbrace{\frac{Q}{4\pi\epsilon_0 r}}_{\text{monopole}} + \underbrace{\frac{\mathbf{p} \cdot \mathbf{r}}{4\pi\epsilon_0 r^3}}_{\text{dipole}} + \underbrace{\frac{1}{8\pi\epsilon_0 r^5} \iiint_{(V)} [3(\mathbf{r}' \cdot \mathbf{r})^2 - r'^2 r^2] \rho(\mathbf{r}') d^3 r'}_{\text{quadrupole}}$$

The expansion is usually truncated at the lowest non-zero term, thereby characterizing the distribution as a monopole $Q \neq 0$, dipole ($Q = 0$ and $\mathbf{p} \neq \mathbf{0}$), quadrupole ($Q = 0$, $\mathbf{p} = \mathbf{0}$ and $D \neq 0$), etc.

- The electrostatic (potential) energy of a charge distribution is

$$U_p = \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{r}) V(\mathbf{r})}{2} d^3 r = \frac{\epsilon_0}{2} \iiint_{\mathbb{R}^3} |\mathbf{E}(\mathbf{r})|^2 d^3 r$$

- The action of an electric field on a dipole distribution consists in a force \mathbf{F} and a torque \mathcal{T} ,

$$\text{force } \mathbf{F} = (\mathbf{p} \cdot \nabla) \mathbf{E}, \quad \text{torque } \mathcal{T} = \mathbf{p} \times \mathbf{E}.$$

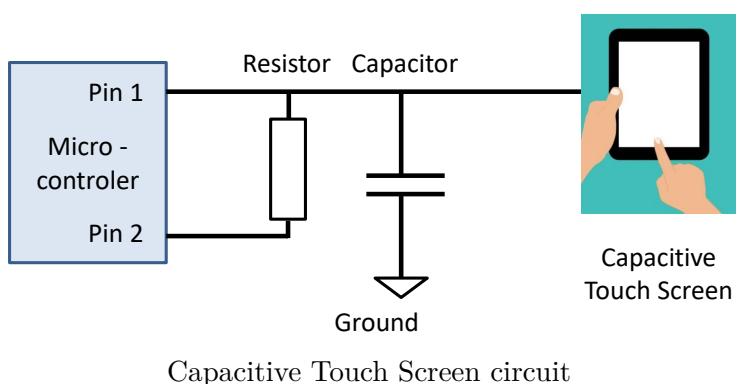
The force is non-zero only for non-uniform electric fields and tend to pull the dipoles toward regions of high-fields. The torque tends to align the dipoles parallel to the fields.

Chapter 3

Metals, conductors and capacitors



The Leyden jar



Introduction

Matter may be roughly classified into two families: conductors and insulators. This chapter deals with the behavior of conductor materials in the presence of electric fields. The study of insulators is left for chapters 4 and 5.

Inside a conductor, a large density of electrons may freely propagate, meaning that these electrons are not attached to any particular atom of the host crystal. In consequence, if an electric field is present inside a conductor, electrons will be accelerated, thus generating a flow of charges known as electric current.

Two possibilities arise:

1. The conductor is isolated: in this case, the current will eventually stop once the displacement of electrons in the medium generates an electric field that cancels the external field, the system reaches a static equilibrium. This case is the specific subject of this chapter.

- The electric field is imposed by an external source: this is the case, for example, when a battery keeps two points of a conductor at a different electric potential. A stationary electric current develops, according to what is known as Ohm's law. This case will be investigated in chapter 6.

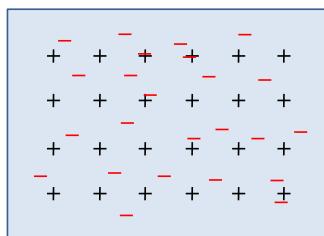
This chapter deals with conductors in static equilibrium. Conductors are used to store electric charges. After establishing a model for metals and perfect conductors, we will review the concepts of capacitance, which measures the capacity of a conductor to store charge. Generalizing this concept to a set of conductors leads to the study of capacitors. The concepts of energy, forces and torques are finally presented for the conductors that constitute a capacitor.

3.1 Model for metals

3.1.1 Fixed and mobile charges

The role of a model for metals is to help understanding their properties and in particular their conductivity, which is due to the motion of electric charges. A microscopic model for the conductivity of metals will be presented in chapter 6. Here, a metal model is presented which allows us to understand the properties of conductors at equilibrium.

A metal, for example copper, is modeled as a lattice of positively charged ionized atoms. These ions only move around their average position in the lattice; this is a small amplitude motion, that is, ions remain localized in the vicinity of their average position and can be assumed as fixed. Their spatially averaged volume charge density ρ_f is constant¹.



Throughout the chapter, we will assume that each atom of the metal liberated one electron, leading to ions of charge $+e$ and free electrons of charge $-e$. There is no loss of generality in this assumption: the same model remains valid if we assume that each atoms liberated n electrons, leading to a globally neutral metal with ions of charge ne and n free electrons per atom, of charge $-e$. Free electrons are mobile in the entire volume of the metal. They are called *charge carriers*. Their spatially averaged charge density is ρ_m ².

¹Subscript f stands for fixed charges

²Subscript m stands for mobile charges

Why do we specify *spatially averaged* for the charge densities? It is possible to view all charges in a metal as point charges and simply express their densities using Dirac delta as $\rho_f = \sum_{i \in \text{ions}} +e\delta(\mathbf{r} - \mathbf{R}_i)$ and $\rho_m = \sum_{i \in \text{electrons}} -e\delta(\mathbf{r} - \mathbf{r}_i(t))$, where the positions of ions \mathbf{R}_i are assumed time independent, and $\mathbf{r}_i(t)$ denotes the positions of free electrons. However, the concept of point charge raises several questions:

- Atoms or ions certainly have a finite size, even if it is small. Can we then still approximate charges as point charges?
- Quantum physics tell us that perfect localization of an electron at a time dependent position $\mathbf{r}_i(t)$ is a purely classical view³. Electrons are described in terms of a wave function from which their probability of presence at a given point in space can be obtained.
- With point charges, all quantities appearing in Maxwell's equations, including charges and current densities, as well as the fields they generate would undergo spatial fluctuations at the microscopic scale (distance between ions in the lattice of a metal), whereas Maxwell's equations were proposed to describe macroscopic situations.

Hendrik Lorentz assumed that Maxwell's equations are valid to describe microscopic situations (point charges) and proposed a spatial averaging scheme to average out fluctuations appearing at the microscopic scale of atoms, leading to smoothly varying charge densities and fields at the macroscopic scale. After the development of *Quantum Electrodynamics* around 1940, it is the agreement between experiments and theory, for instance for the atomic spectrum of hydrogen, that confirmed the validity of Lorentz's assumption that Maxwell's equations applies in the realm of the microscopic world. In classical electrodynamics, we will focus on macroscopic situations. The next sections provide details on Lorentz's procedure of spatial averaging. In the rest of this chapter, both ρ_m and ρ_f will denote spatially averaged charge densities.

3.1.2 Average density of carriers

Consider a volume element d^3r around point M , at time t . This volume contains a total charge dq_f for the fixed ions and dq_m for mobile charges. The typical distance between two ions in the lattice of the metal is 10^{-10} m. Assume the volume element has a typical size of 10^{-6} m in each dimension. Hence, it contains $\sim 10^{12}$ charges⁴. This quantity is sufficient to define an average, yet the average charge densities ρ_f and ρ_m (in m^{-3}) defined by

$$dq_f = \rho_f(\mathbf{r}, t)d^3r, \quad dq_m = \rho_m(\mathbf{r}, t)d^3r,$$

vary smoothly at the scale of one micron (10^{-6} m). The scale used to define a spatial average, often called a *mesoscopic* scale⁵ does not need to be exactly one micron. It is sufficient that it

³It would imply that, e.g., the electron velocity $\mathbf{v}_i(t)$ is simultaneously known in violation of the Heisenberg inequality.

⁴This is an order of magnitude, assuming a compact arrangement of 10^4 spheres in each dimension.

⁵It is much larger than the microscopic scale and much smaller than the macroscopic scale.

is large enough compared to the macroscopic scale of the problem and small enough compared to the microscopic scale (Bohr's radius or scale of an atom) for the spatial average to make sense.

3.1.3 Spatially averaged electric fields

Maxwell's equations deal with macroscopic quantities (electric and magnetic fields, charge and current densities). It is Hendrik Lorentz who proposed that Maxwell's equations apply to microscopic situations too, provided the source densities $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ are microscopic, that is, vary rapidly on a microscopic scale. Lorentz proposed a derivation of Maxwell's macroscopic equations from their microscopic counterpart, using a spatial averaging procedure leading to slowly varying macroscopic sources.

Consider an averaging region $\Omega_{\mathbf{r}}$ around each point \mathbf{r} , such that its volume $\mathcal{V}(\Omega_{\mathbf{r}})$ is much larger than the typical volume a_B^3 of a cube of side equal to the Bohr radius a_B , $\Omega \gg a_B^3$. The exact boundary or shape of $\Omega_{\mathbf{r}}$ is not important as long as it contains at least 10 to 100 particles in each dimension, that is 10^3 to 10^6 particles in three dimensions or equivalently $(10a_B)^3 < \Omega \leq (100a_B)^3$. Equivalently, the volume $\mathcal{V}(\Omega_{\mathbf{r}})$ can be chosen constant⁶ and equal to the inverse of the average density of atoms $\mathcal{V} \simeq n^{-1}$ so as to fulfill the condition on the number of particles⁷. This volume is called a *mesoscopic* volume.

If $\mathbf{e}(\mathbf{r}', t)$ denotes the electric field at each microscopic point \mathbf{r}' , the spatially averaged electric field then reads

$$\mathbf{E}(\mathbf{r}, t) = \langle \mathbf{e}(\mathbf{r}, t) \rangle = \iiint_{\mathbb{R}^3} \mathbf{e}(\mathbf{r}', t) \phi(\mathbf{r} - \mathbf{r}') d^3 r' = \frac{1}{\mathcal{V}} \iiint_{\Omega_{\mathbf{r}}} \mathbf{e}(\mathbf{r} + \mathbf{r}'', t) d^3 r'',$$

where the function $\phi(\mathbf{r})$ takes the value $1/\mathcal{V}$ if $\mathbf{r} \in \Omega_0$ and zero otherwise. and varies slowly on the scale of the variable \mathbf{r} . The typical scale over which the macroscopic quantities vary is thus equal to 10 to 100 a_B , or equivalently to $n^{-1/3}$. The macroscopic electric field $\mathbf{E}(\mathbf{r}, t)$ that enters in Maxwell's equations corresponds to the spatially and averaged quantity $\langle \mathbf{e}(\mathbf{r}, t) \rangle$. All other quantities appearing in Maxwell's equations correspond similarly to the spatial average of their microscopic counterpart. This averaging procedure amounts to considering point charges, for instance electrons in a metal, as a fluid (or a gas) of electric charges. With this model in mind, we will see how to determine the conductivity of metals.

3.2 The Drude model for electric conductivity

In the beginning of the 20th century (1900), the German physicist Paul Drude constructed a simple model capable of describing qualitatively the conductivity of metals by considering electrons as a gas of classical particles. In the absence of an electric field, the velocity \mathbf{v}_i of the

⁶only its center changes with \mathbf{r} .

⁷In the SI unit system, n is expressed in m^{-3} .

i^{th} electron is randomly distributed and when averaged over timescales much longer than the time between two collisions we obtain $\langle \mathbf{v}_i \rangle = \mathbf{0}$. However, under the influence of a constant electric field \mathbf{E} applied at $t = 0$, an electron of velocity \mathbf{v}_i , mass m_e and charge $-e$ will be accelerated according to Newton's second law:

$$m_e \frac{d\mathbf{v}_i}{dt} = -e\mathbf{E},$$

so that it acquires an increasing velocity in the direction of the electric field

$$\mathbf{v}_i(t) = \mathbf{v}_i(0) - \frac{et}{m_e} \mathbf{E}.$$

However, it is known from experience that a steady current develops in a conductor whenever an electric field is applied (Ohm's law, which will be studied in chapter 6). This means than electrons cannot accelerate indefinitely, they must acquire a finite steady state velocity. Of course, the equation above is only valid between two collisions, since an electron will give away part of its kinetic energy during each collision. Hence, the finite steady state velocity corresponds to a temporally averaged velocity over a duration much longer than the duration between two successive collisions. The effect of collisions may be taken into account by including a friction term in the equation of motion, describing the dissipation of energy under the form of heat (this is called the Joule effect - see chapter 6). Let us then write the equation of motion for the average velocity $\langle \mathbf{v}_i \rangle$ in which, in addition to the electric force, we add a friction term proportional to the velocity

$$m_e \frac{d\langle \mathbf{v}_i \rangle}{dt} = -e\mathbf{E} - m_e \frac{\langle \mathbf{v}_i \rangle}{\tau_e}, \quad (3.1)$$

where τ_e has units of time will be shown in chapter 6 to represent an average collision time. Equation (3.1) admits a steady state solution for the velocity

$$\langle \mathbf{v}_i \rangle = -\frac{e\tau_e}{m_e} \mathbf{E} = \mu_e \mathbf{E}, \quad (3.2)$$

where $\mu_e = -e\tau_e/m_e$ is the electron **mobility**, which in a given material determines the proportionality constant between the external electric field and the steady state average velocity of the electrons.

In this section, we have calculated a temporally average velocity over a time scale much larger than the collision time τ_e , and we used the same bracket notation $\langle \rangle$ for the average quantity as that used in previous section for Lorentz's spatially averaged fields. From Statistical Physics, we learn that under conditions of equilibrium which are satisfied here, we can equivalently average in time, or average spatially (that is over the electrons contained within a mesoscopic volume). Hence, considering the electrons as a fluid of electricity with (spatially averaged) velocity $\mathbf{v}(\mathbf{r}, t)$, we can write the equation of state for a metal as

$$\mathbf{v} = \mu_e \mathbf{E}. \quad (3.3)$$

3.3 Conductors in static equilibrium

In this section we will review the main properties of conductors in electrostatic equilibrium, which constrains how the electric field and the electric potential vary in space.

3.3.1 The electric field in a conductor at equilibrium is zero

Since an electron will be accelerated as soon as a non-zero electric field is present, eventually reaching a steady state average velocity satisfying the equation of state for a metal $\mathbf{v} = \mu_e \mathbf{E}$, the only acceptable solution for static equilibrium is obtained when the electric field is zero at every point inside a conductor. If Ω represents the volume of a conductive material, then

$$\mathbf{E}(\mathbf{r}) = \mathbf{0} \quad \forall \mathbf{r} \in \Omega$$

which can be written in terms of the potential as

$$-\nabla V(\mathbf{r}) = 0 \quad \forall \mathbf{r} \in \Omega$$

so that the electrostatic potential is uniform inside a conductor in static equilibrium.

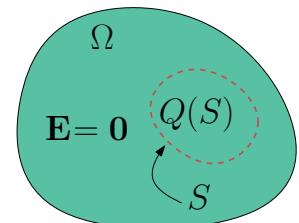
3.3.2 The charge of a conductor in equilibrium is distributed on its surface

A consequence of the previous property is that every charge imbalance in a conductor Ω must reside on its surface $\partial\Omega$. To show this, imagine an arbitrary closed surface $S \in \Omega$. Since the electric field is zero in Ω , the electric flux through S is zero, and according to Gauss's law this implies that the charge $Q(S)$ contained in the volume inside S must be zero:

$$\Phi_{S,\mathbf{E}} = \iint_S \mathbf{E} \cdot \mathbf{n} dS = \frac{Q(S)}{\epsilon_0} = 0.$$

Since S is an arbitrary surface contained in Ω , it follows that the charge density ρ is zero everywhere inside a conductor

$$\rho(\mathbf{r}) = 0 \quad \forall \mathbf{r} \in \Omega$$



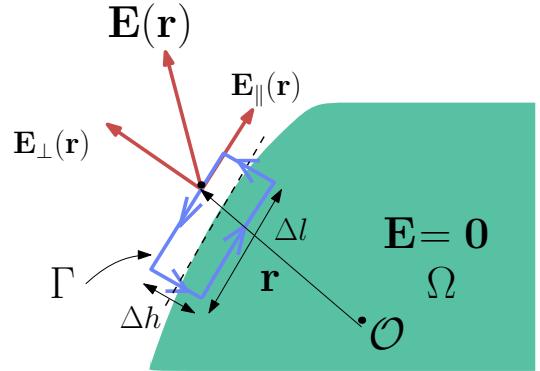
and consequently, in a conductor at equilibrium, all the charge must be distributed on its surface. A conductor in equilibrium is thus completely characterized by its surface charge density σ in $\partial\Omega$.

3.3.3 The electric field is normal to the surface of a conductor in equilibrium

For the electrons on the surface of a conductor to be at rest, the electric field must be normal to the surface at every point. Otherwise, a non-zero tangential component will put the electrons of the surface in movement. In fact, since the electric field is null in Ω , it is necessarily perpendicular to the surface. To show this, let us consider a point \mathbf{r} on the surface ($\mathbf{r} \in \partial\Omega$) and decompose the electric field into the sum of its perpendicular components, normal and parallel to the surface, $\mathbf{E}(\mathbf{r}) = \mathbf{E}_\perp(\mathbf{r}) + \mathbf{E}_\parallel(\mathbf{r})$. We define the infinitesimal closed planar path Γ around \mathbf{r} , as shown in the figure below.

The circulation law for the electric field writes

$$\oint_{\Gamma} \mathbf{E} \cdot d\mathbf{l} = 0.$$



The dimensions of the curve Γ being infinitesimal, we neglect the spatial variation of the electric field along the upper and lower parts of the path. Then, the circulation along the two segments parallel to the normal \mathbf{n} cancel each other, and we are left with

$$\oint_{\Gamma} \mathbf{E} \cdot d\mathbf{l} \approx -|\mathbf{E}_\parallel(\mathbf{r})|\Delta l = 0$$

which means that, for every $\mathbf{r} \in \partial\Omega$, $\mathbf{E}_\parallel(\mathbf{r}) = 0$ and the electric field is of the form

$$\boxed{\mathbf{E}(\mathbf{r}) = E(\mathbf{r})\mathbf{n}(\mathbf{r}) \quad \forall \mathbf{r} \in \partial\Omega}.$$

This also means that the surface of a conductor is an equipotential, since for any pair of points A and B on the surface, we have

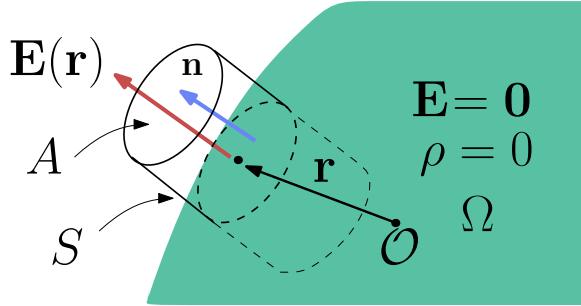
$$V(A) - V(B) = \int_A^B \mathbf{E} \cdot d\mathbf{l}$$

and choosing a path contained in $\partial\Omega$

$$V(A) - V(B) = \int_A^B \underbrace{\mathbf{E} \cdot d\mathbf{l}}_{\mathbf{E}_\parallel \cdot d\mathbf{l} = 0} = 0$$

3.3.4 Coulomb's theorem: electric field at the surface of a conductor in equilibrium

Finally, let us review the relationship between the electric field \mathbf{E} and the charge density σ at the surface of a conductor in equilibrium. For this, we apply Gauss's law to an arbitrarily-small closed cylindrical surface S around $\mathbf{r} \in \partial\Omega$



The electric flux is non-zero only through the top surface, that we suppose to have an area A , and the charge contained in the volume enclosed by S is the charge σA present on the conductor surface $\partial\Omega$ intersected by the cylinder. Gauss's law then reads

$$\iint_S \mathbf{E} \cdot \mathbf{n} dS = E(\mathbf{r})A + 0A = \frac{\sigma(\mathbf{r})A}{\epsilon_0}$$

and so we obtain Coulomb's theorem

$$\mathbf{E}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\epsilon_0} \mathbf{n}, \quad \mathbf{r} \in \partial\Omega.$$

(3.4)

Remark

A charged surface with density σ generates an electric field above (at M) and below the surface (at M') which is given by $\mathbf{E} = \pm\sigma/(2\epsilon_0)\mathbf{n}$, respectively. Consider now a conductor and point M (resp. M') located at infinitesimal distance above (resp. below) the conductor surface, close enough to the surface to be considered as belonging to the surface for applying Coulomb's theorem. From the point of view of M the infinitesimal surface dS around M carrying the charge σdS , behaves like an infinite plane, hence, it generates a field $\mathbf{E}_{dS}(M) = \sigma/(2\epsilon_0)\mathbf{n}$ outside the conductor and a field $\mathbf{E}_{dS}(M') = -\sigma/(2\epsilon_0)\mathbf{n}$ inside. If \mathcal{R} denotes the rest of the conductor surface, from the superposition theorem and Coulomb's theorem,

$$\mathbf{E}(M) = \mathbf{E}_{dS}(M) + \mathbf{E}_{\mathcal{R}}(M) = \frac{\sigma}{\epsilon_0} \mathbf{n} \text{ above the surface,}$$

and

$$\mathbf{E}(M') = \mathbf{E}_{dS}(M') + \mathbf{E}_{\mathcal{R}}(M') = \mathbf{0} \text{ below the surface.}$$

We infer that \mathcal{R} generates an electric field $\mathbf{E}_{\mathcal{R}}(M') = \sigma/(2\epsilon_0)\mathbf{n}$ in the region inside the conductor that exactly cancels out the field generated by the charges on dS . In the external region, the field generated by the surface dS and that generated by \mathcal{R} are equal: $\mathbf{E}_{dS}(M) = \mathbf{E}_{\mathcal{R}}(M) = \frac{\sigma}{2\epsilon_0}\mathbf{n}$ and add up to give a total field $\frac{\sigma}{\epsilon_0}\mathbf{n}$.

3.3.5 Electrostatic energy of a conductor

Since all the charge in a conductor Ω is distributed on its surface $\partial\Omega$, and since the potential V is constant in $\partial\Omega$, the electrostatic energy of the conductor writes

$$U = \frac{1}{2} \iint_{\partial\Omega} V(\mathbf{r}') dq(\mathbf{r}') = \frac{1}{2} V_0 \underbrace{\iint_{\partial\Omega} \sigma(\mathbf{r}') dS(\mathbf{r}')}_Q,$$

where V_0 is the potential in the conductor and we recognize Q , the total charge on its surface, in the integral on the right-hand side. Finally

$$U = \frac{1}{2} V_0 Q. \quad (3.5)$$

3.4 Interaction between conductors

When two or more conductors touch each other, and if at least one of them is charged, the charges at their surfaces must be distributed in such a way that the electric field is zero inside every conductor and, according to Coulomb's theorem, equal to $\mathbf{E}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\epsilon_0}\mathbf{n}(\mathbf{r})$, where $\mathbf{n}(\mathbf{r})$ is the normal to a conductor surface at point \mathbf{r} . This is the main reason why when dealing with conductors other mathematical tools than the Coulomb integral⁸ are necessary to determine the electric field: the charge densities are not known a priori.

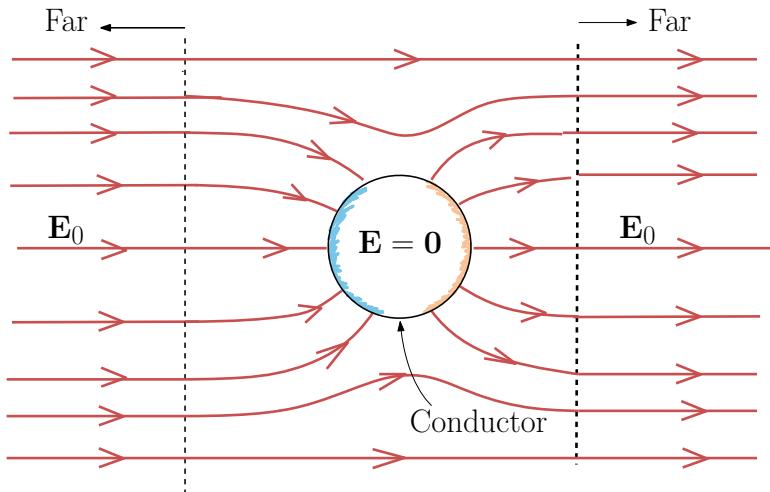
3.4.1 Electrostatic induction

Let us suppose that a neutral conductor is placed in a region in which there is an electric field \mathbf{E}_0 . At equilibrium, the electric field must vanish inside the conductor. For this to happen, a

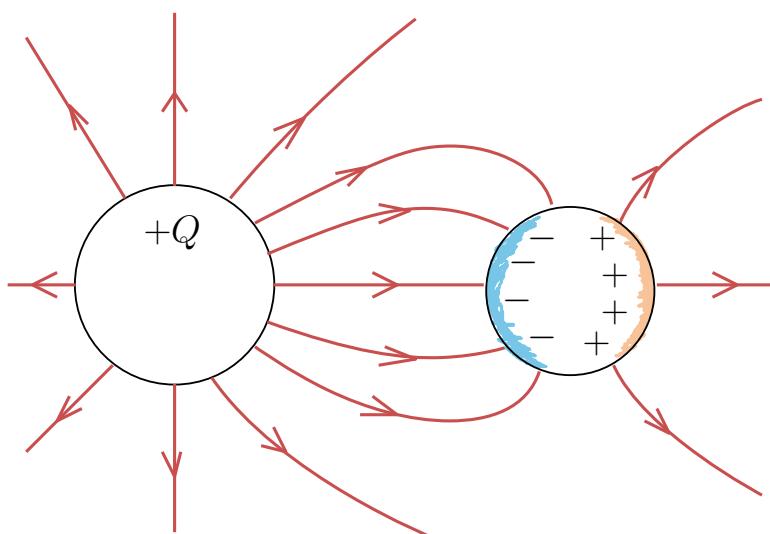
⁸Coulomb's integral:

$$\mathbf{E}(\mathbf{r}) = \iiint_{\Omega} \frac{1}{4\pi\epsilon_0} \frac{\rho(\mathbf{r}')(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} d^3 r'$$

charge separation must occur at the surface of the conductor in order to produce an electric field capable of canceling the external field inside the conductor. The conductor is then polarized, i.e., negative charges accumulate on one side, leaving a positive charge of equal magnitude on the opposite side. This redistribution of charges in a conductor is called electrostatic induction.

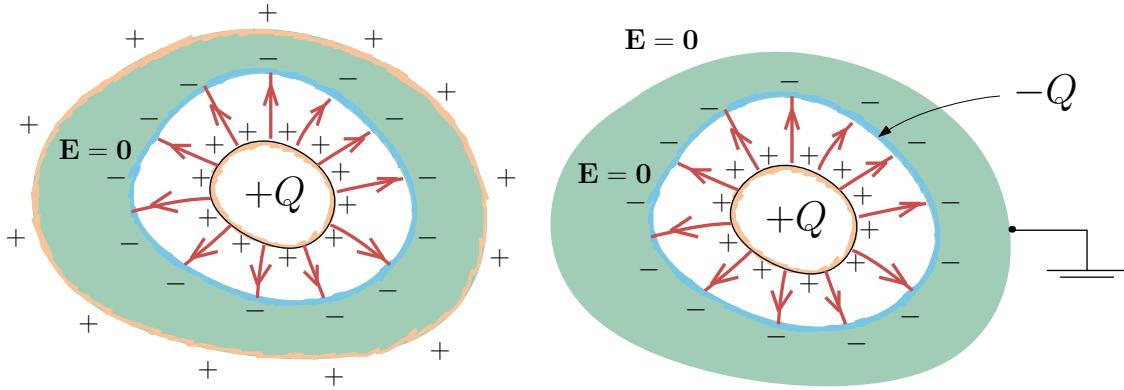


This phenomenon of induction or *influence* happens every time two or more conductors get close to each other. The redistribution of charges in a conductor can be easily interpreted as an attraction of negative charges towards the positive charges on the other conductors, leaving an equal amount of positive charge on the opposite side.



We say that the induction between two conductors is total if all the field lines start from one conductor and end on the other. This can be achieved if, for example, one conductor is completely surrounded by another one. In the figure below, a conductor with charge Q is surrounded by a conductive shell. The charges at the surfaces of the shell rearrange so as to cancel the electric field at any point inside it. This means that the induced charge in the inner surface is $-Q$ (easy to show with Gauss's law), thus leaving a charge $+Q$ on the outer surface.

If the latter is connected to the ground, this positive charge is evacuated so that the potential of the shell is equal to the potential of the ground. The shell gets a total charge $-Q$, and all the field lines connect the two conductors, which are therefore under total influence.



3.4.2 Capacitance

- CAPACITANCE OF AN ISOLATED CONDUCTOR

The capacitance of a conductor in electrostatic equilibrium measures the capacity of a conductor to store charge. It is defined from the relation existing between the electric potential V of the conductor and the charge Q on its surface. Let Ω be the volume occupied by a conductor, then

$$Q = \iint_{\partial\Omega} \sigma(\mathbf{r}') dS(\mathbf{r}'),$$

$$V = \frac{1}{4\pi\epsilon_0} \iint_{\partial\Omega} \frac{\sigma(\mathbf{r}') dS(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \text{ for } \mathbf{r} \in \Omega.$$

We see that if the charge density σ is scaled as $\sigma' = \alpha\sigma$, so does the charge and the potential:

$$Q' = \alpha Q \quad V' = \alpha V$$

In consequence, $Q/V = Q'/V'$ is a constant depending solely on the geometry of the conductor Ω .

Definition 3.1: Capacitance of a conductor

The capacitance of a conductor in equilibrium is a constant that depends on its geometry, and defined as

$$C = Q/V, \tag{3.6}$$

where V is the absolute potential of the conductor when the latter has a total charge Q on its surface. In S.I. units, the unit for the capacitance is the Farad (F), with $1 \text{ F} = 1 \text{ C V}^{-1}$. In electronics, capacitances are rarely larger than one microFarad ($1 \mu\text{F} = 1 \times 10^{-6} \text{ F}$).

- CAPACITANCE AND ELECTROSTATIC ENERGY

In section 3.3.5 the electrostatic energy of a charged conductor was shown to be

$$U = \frac{1}{2} QV.$$

In terms of the capacity ($Q = CV$), this can be rewritten as

$$U = \frac{1}{2} CV^2 = \frac{Q^2}{2C}. \quad (3.7)$$

3.4.3 Generalization: induction coefficient between conductors

Suppose a system of N conductors in equilibrium, each one with charge Q_i and potential V_i , with $i \in \{1, 2, \dots, N\}$. Let each conductor be represented by a macroscopic volume Ω_i with a surface $\partial\Omega_i$, carrying a surface charge density σ_i . The charges Q_i and the potentials V_i of the conductors read

$$Q_i = \iint_{\partial\Omega_i} \sigma_i(\mathbf{r}) dS(\mathbf{r}),$$

$$V_i = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^N \iint_{\partial\Omega_j} \frac{\sigma_j(\mathbf{r}) dS(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} \text{ for } \mathbf{r}_i \in \Omega_i.$$

where $\mathbf{r}_i \in \Omega_i$ and the integral for the potentials is performed over the total surface of all conductors. This shows clearly the linear relationship between charges and potentials: if any of the charge densities $\sigma_i(\mathbf{r})$ is scaled as $\alpha\sigma_i(\mathbf{r})$, the charge Q_i on conductor i is also multiplied by alpha and the term corresponding to $j = i$ in the potential V_i will be multiplied by α as well. As a consequence, we can write

$$Q_i = \sum_{j=1}^N C_{ij} V_j.$$

or as a matrix by vector product

$$\begin{array}{c|c} Q_1 & \left[\begin{matrix} C_{11} & \dots & C_{1N} \end{matrix} \right] | V_1 \\ \vdots & \left[\begin{matrix} \vdots & \ddots & \vdots \end{matrix} \right] | \vdots \\ Q_N & \left[\begin{matrix} C_{N1} & \dots & C_{NN} \end{matrix} \right] | V_N \end{array}$$

Definition 3.2: Induction coefficient

The induction coefficient between the conductors i and j is defined as

$$C_{ij} = \frac{\partial Q_i}{\partial V_j}. \quad (3.8)$$

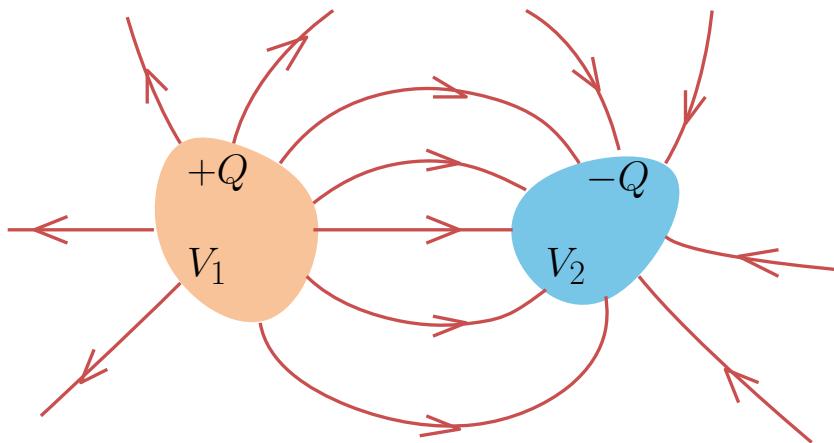
Now, if a charge Q_i on conductor i brings the conductor j to a potential V_j , then by symmetry the same charge Q_i on conductor j would bring the conductor i to a potential V_j . It results that $C_{ij} = C_{ji}$. Generalizing the case of an isolated conductor, the self-capacitance coefficient of the conductor i under the presence of the $N-1$ other conductors is defined as

$$C_{ii} = \frac{\partial Q_i}{\partial V_i} \quad (3.9)$$

whereas the C_{ij} coefficients with $i \neq j$ are called the mutual capacitance coefficients.

3.5 The capacitor

A capacitor is a system of two conductors carrying charges of equal magnitude but of opposite sign, as shown in the figure below.



This system is globally neutral, so that far away it generates an electric field equivalent to that of an electric dipole. A significant fraction of the field lines is then confined in the region between the two conductors, and so does the electrostatic energy density of the system.

3.5.1 Capacitance of a capacitor

Suppose that a capacitor is made of two conductors carrying charges $Q_1 = Q$ and $Q_2 = -Q$, at potentials V_1 and V_2 , respectively. We have

$$\begin{aligned}Q_1 &= +Q = C_{11}V_1 + C_{12}V_2, \\Q_2 &= -Q = C_{21}V_1 + C_{22}V_2,\end{aligned}$$

and since $C_{12} = C_{21}$, this can be rewritten as

$$\begin{aligned}Q &= C_{11}V_1 + C_{12}V_2, \\-Q &= C_{12}V_1 + C_{22}V_2.\end{aligned}$$

The sum of these two equations gives

$$0 = (C_{12} + C_{11})V_1 + (C_{12} + C_{22})V_2.$$

Since the potentials can be shifted arbitrarily (one can always add an arbitrary constant to the potentials without changing the electric field), this implies $C_{11} = C_{22} = -C_{12}$ as the above equation must be satisfied whether we subtract V_1 or V_2 to the potentials⁹. We therefore find a linear relationship between the charge Q and the potential difference between the conductors.

$$Q = C(V_1 - V_2) = C\Delta V, \quad (3.10)$$

where C is, by definition, the capacitance of the capacitor,

$$C = \frac{Q}{\Delta V} \quad (3.11)$$

and represents its capacity to produce a charge separation for a given potential difference. We see that a set of two conductors is characterized in general by the capacitance matrix that contains three coefficients: the two self-capacitance- and the mutual capacitance-coefficients. Only when the two conductors carry opposite charges, that is when they form a capacitor, do these three coefficients reduce to a single capacitance coefficient characterizing the capacitor.

3.5.2 Electrostatic energy of a capacitor

For a general system of conductors, we have, from Eq. (3.5), a total electrostatic energy of

$$U = \frac{1}{2} \sum_i Q_i V_i.$$

⁹If we subtract V_1 , then $(C_{12} + C_{22})(V_2 - V_1) = 0$, ie $C_{12} + C_{22} = 0$. If we subtract V_2 , then $(C_{12} + C_{11})(V_1 - V_2) = 0$, ie, $C_{12} + C_{11} = 0$.

In the case of a capacitor, we then find

$$U = \frac{1}{2}[QV_1 - QV_2] = \frac{1}{2}Q\underbrace{(V_1 - V_2)}_{\Delta V}$$

and the electrostatic energy contained in the capacitor is

$$U = \frac{1}{2}Q\Delta V = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}C\Delta V^2. \quad (3.12)$$

3.5.3 Determination of a capacitance

In order to determine the capacitance of a capacitor, the following general procedure can be applied:

1. Assume a charge Q on one of the conductors, and therefore a charge $-Q$ on the other one.
2. The electric field \mathbf{E} between the conductors is then determined.
3. The potential difference is given by

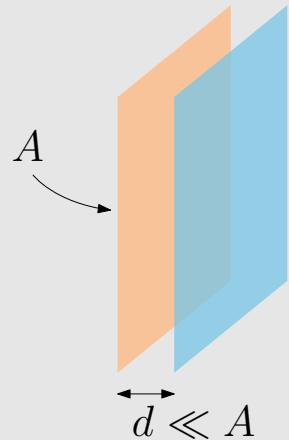
$$\Delta V = V_1 - V_2 = \int_1^2 \mathbf{E} \cdot d\mathbf{l}.$$

The most appropriate path to calculate this line integral will depend on the symmetries of the system.

4. Finally, the capacitance is given by $C = Q/\Delta V$, which is independent on the charge Q .

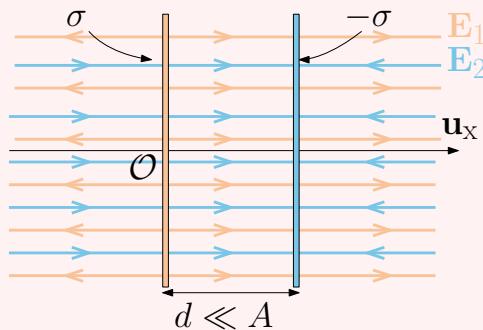
Example 3.1 - The parallel plate capacitor

Consider a capacitor made of two parallel plates of area A separated by a distance $d \ll A$. Determine the capacitance by neglecting edge effects, that is, by considering that both plates are of infinite extension (this is a good approximation at any point between the plates away from the edges since $d \ll A$).



Solution

Suppose that both plates are homogeneously charged with surface densities σ and $-\sigma$, respectively. Let Ox be the axis perpendicular to both plates, and let us choose the origin \mathcal{O} at the center of the left plate, as shown in figure below.



The electric field will be the superposition of the fields generated by each plane separately. Recalling the electric field generated by an infinite plane carrying a surface charge, we have

$$\begin{aligned}\mathbf{E}_1(\mathbf{r}) &= \frac{\sigma}{2\epsilon_0} \frac{x}{|x|} \mathbf{u}_x, \\ \mathbf{E}_2(\mathbf{r}) &= \frac{-\sigma}{2\epsilon_0} \frac{x-d}{|x-d|} \mathbf{u}_x.\end{aligned}$$

The superposition of both fields vanishes everywhere except for the region between the plates, so that the plates are at total influence and the field between them is given by

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_1(\mathbf{r}) + \mathbf{E}_2(\mathbf{r}) = \begin{cases} \frac{\sigma}{\epsilon_0} \mathbf{u}_x & \text{if } 0 < x < d, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

The field lines are thus confined in the region between the plates and so does the electrostatic energy density. The potential difference between the plates is

$$\Delta V = V_1 - V_2 = \int_0^d \mathbf{E} \cdot d\mathbf{r}$$

By choosing a path parallel to the Ox axis,

$$\Delta V = \int_0^d \frac{\sigma}{\epsilon_0} dx = \frac{\sigma d}{\epsilon_0}$$

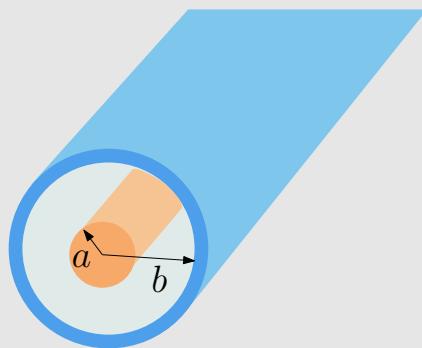
If the area of each plate is A , then the total charge on the left plate is given by $Q = A\sigma$ and the capacity reads

$$C = \frac{Q}{\Delta V} = \frac{\epsilon_0 A}{d}.$$

For a capacitor made of square plates of lateral size of 10 cm, we have $A = 1 \times 10^{-2} \text{ m}^2$, and supposing a separation of $d = 1 \text{ cm}$, the capacity is $C = 9 \times 10^{-12} \text{ F}$. If a potential difference of $\Delta V = 1.5 \text{ V}$ is applied, the charge separation between the plates will be $Q = C\Delta V = 1.5 \times 9 \times 10^{-12} = 1.35 \times 10^{-11} \text{ C}$ and the stored electrostatic energy is $U = \frac{1}{2}Q\Delta V = 1.01 \times 10^{-12} \text{ J}$.

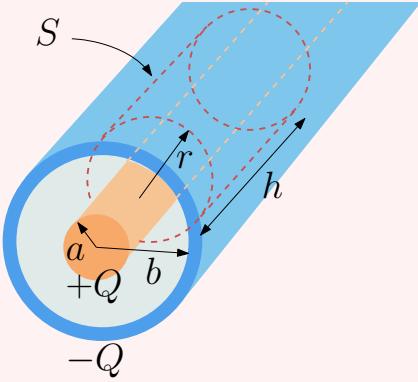
Example 3.2 - Cylindrical capacitor

Consider a capacitor made of two cylindrical and concentric conductors, of radius a and b , and length $L \gg (b - a)$. By neglecting edge effects, determine the capacity of this capacitor.



Solution

The capacitance may be obtained by supposing that the inner conductor carries a charge Q and the outer one $-Q$. Neglecting edge effects, the charge distribution has a cylindrical symmetry and the electric field writes, in cylindrical coordinates, $\mathbf{E}(\mathbf{r}) = E(r)\mathbf{u}_r$.



Choosing as Gauss closed surface S a cylinder of radius r , with $a < r < b$, concentric to the conductors and of length h , we have, according to Gauss's law

$$\iint_S \mathbf{E}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) dS(\mathbf{r}) = E(r)2\pi rh = \frac{Q(S)}{\epsilon_0}.$$

The charge inside the surface S is $\frac{Q}{L}$, and so

$$E(r)2\pi rh = \frac{Qh}{\epsilon_0 L}.$$

Hence, the electric field writes

$$\mathbf{E}(\mathbf{r}) = \frac{Q}{2\pi\epsilon_0 r L} \mathbf{u}_r.$$

Note that for $r < a$ and $r > b$ the charge $Q(S)$ inside S is zero; the electric field is confined in between the conductors. The potential difference is

$$V(b) - V(a) = - \int_a^b \mathbf{E} \cdot d\mathbf{l} = - \int_a^b E(r) dr,$$

where a radial path has been used ($d\mathbf{l} = dr \mathbf{u}_r$). Then

$$\Delta V = \frac{Q}{2\pi L \epsilon_0} \int_a^b \frac{dr}{r} = \frac{Q}{2\pi L \epsilon_0} \ln\left(\frac{b}{a}\right)$$

Finally, the capacitance is

$$C = \frac{Q}{\Delta V} = \frac{2\pi\epsilon_0 L}{\ln\left(\frac{b}{a}\right)}.$$

Note that in the limit $d = b - a \ll a, b$, we should retrieve the result obtained for the parallel plate capacitor (Ex. 3.1). Indeed, for $d \ll a$

$$\log(b/a) = \log(1 + d/a) \approx d/a$$

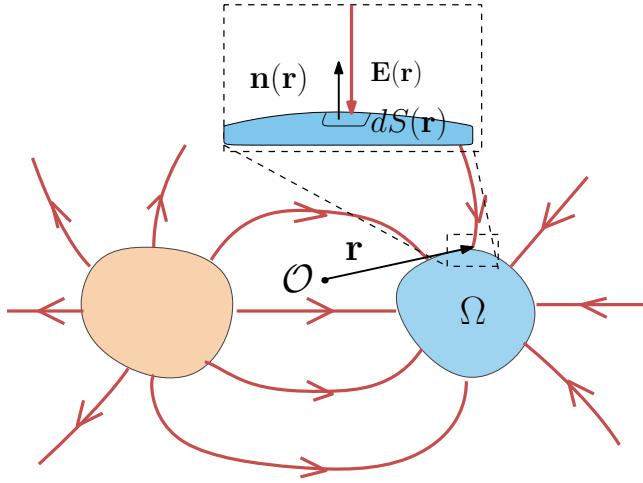
and so,

$$C \approx \frac{2\pi\epsilon_0 a L}{d} = \frac{\epsilon_0 A}{d}$$

with $A = 2\pi a L$ the area of the inner cylinder.

3.6 Force on a charged conductor: the electrostatic pressure

Consider a conductor Ω in electrostatic equilibrium in the presence of an arbitrary distribution of charges. At any point \mathbf{r} outside Ω yet infinitely close to its surface, there is an electric field normal to $\partial\Omega$.



Therefore, a force is exerted on a differential surface element $dS(\mathbf{r})$ around \mathbf{r} since it contains a charge $dq(\mathbf{r}) = \sigma(\mathbf{r})dS(\mathbf{r})$. However, the total electric field at \mathbf{r} is due to both the surface element $dS(\mathbf{r})$ itself and the rest of the system. We can then write

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_{dS}(\mathbf{r}) + \mathbf{E}_{ext}(\mathbf{r}),$$

where $\mathbf{E}_{dS}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{2\epsilon_0}\mathbf{n}(\mathbf{r})$ is the electric field due to the charge in $dS(\mathbf{r})$ (identical to the field generated by a charged infinite plane), while $\mathbf{E}_{ext}(\mathbf{r})$ is the field generated by all the other charges present in the system. By Coulomb's theorem, we have

$$\mathbf{E}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\epsilon_0}\mathbf{n}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{2\epsilon_0}\mathbf{n}(\mathbf{r}) + \mathbf{E}_{ext}(\mathbf{r})$$

so that the external field acting on $dS(\mathbf{r})$ writes

$$\mathbf{E}_{ext}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{2\epsilon_0}\mathbf{n}(\mathbf{r}). \quad (3.13)$$

The electrostatic force acting on the charge $dq(\mathbf{r})$ contained in $dS(\mathbf{r})$ is therefore

$$d\mathbf{F}_e = \underbrace{dq(\mathbf{r})}_{\sigma(\mathbf{r})dS(\mathbf{r})} \mathbf{E}_{ext}(\mathbf{r}) = \frac{\sigma^2(\mathbf{r})}{2\epsilon_0}dS(\mathbf{r})\mathbf{n}(\mathbf{r}).$$

Definition 3.3: The electrostatic pressure

An electrostatic force $d\mathbf{F}_e$ acts on any differential surface element dS of a charged conductor Ω . For $\mathbf{r} \in \partial\Omega$, this force is given by

$$d\mathbf{F}_e(\mathbf{r}) = P_e(\mathbf{r})dS(\mathbf{r})\mathbf{n}(\mathbf{r}),$$

where P_e is the electrostatic pressure

$$P_e(\mathbf{r}) = \frac{\sigma^2(\mathbf{r})}{2\epsilon_0}. \quad (3.14)$$

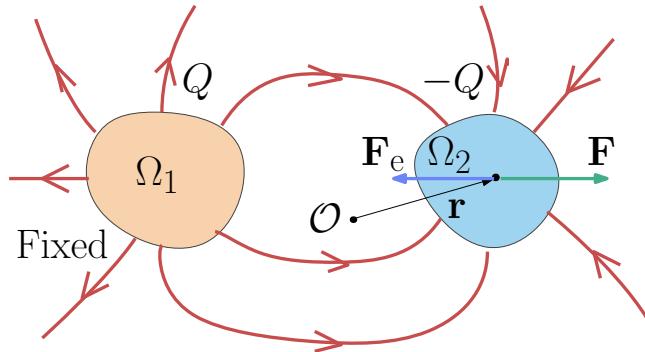
The total electrostatic force on the conductor is therefore given by

$$\mathbf{F}_e = \iint_{\partial\Omega} P_e(\mathbf{r})\mathbf{n}(\mathbf{r})dS(\mathbf{r}) = \iint_{\partial\Omega} \frac{\sigma^2(\mathbf{r})}{2\epsilon_0}\mathbf{n}(\mathbf{r})dS(\mathbf{r}).$$

An external agent must then provide a force that cancels the electrostatic force in order to keep the charged conductors at fixed positions.

3.6.1 Force on a capacitor

In the case of a capacitor, a force must be provided in order to cancel the electrostatic attraction \mathbf{F}_e existing between the two conductors of opposite charge. Consider for example the case of a capacitor made of a fixed conductor Ω_1 and a mobile conductor Ω_2 , whose center of mass is located at position \mathbf{r} .



According to Eq. (3.12), the electrostatic energy of the system writes

$$U = \frac{1}{2}Q\Delta V = \frac{1}{2C}Q^2,$$

where C is the capacitance. This energy will depend on the position \mathbf{r} of the mobile conductor. The electrostatic force \mathbf{F}_e between the conductors is then given by $\mathbf{F}_e = -\nabla U(\mathbf{r})$.

Isolated system, fixed charge

If the system is isolated, there is no possible charge transfer between the conductors, so that Q is a constant independent on \mathbf{r} and

$$U(\mathbf{r}) = \frac{1}{2}Q\Delta V(\mathbf{r}) = \frac{1}{2C(\mathbf{r})}Q^2$$

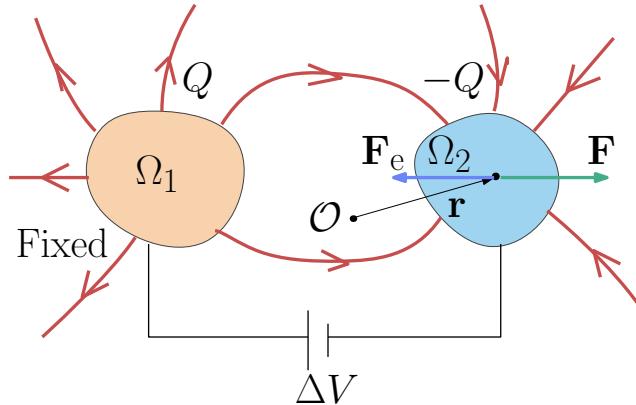
and in this case the external force must balance the electrostatic force, so that $\mathbf{F} = -\mathbf{F}_e$ and

$$\mathbf{F} = (\nabla U)_Q = \frac{Q}{2}\nabla(\Delta V(\mathbf{r})) = -\frac{Q^2}{2C(\mathbf{r})^2}\nabla C(\mathbf{r}).$$

Fixed potential difference

Now consider a more subtle case in which the potential difference between the conductors is kept constant. This can be achieved if the capacitor is connected to a battery or a voltage source imposing a difference ΔV between the two conductors.

If the mobile conductor is displaced by $d\mathbf{r}$, a charge transfer between the conductors must occur in order to keep ΔV fixed. Indeed, we have $\nabla Q(\mathbf{r}) = \Delta V \nabla C(\mathbf{r})$. This charge redistribution costs an additional energy $W = dQ \Delta V = \Delta V \nabla Q(\mathbf{r}) \cdot d\mathbf{r}$, which must be provided (if $dQ > 0$) or consumed (if $dQ < 0$) by the battery.



The work of the external force is then

$$\mathbf{F} \cdot d\mathbf{r} = -W + \underbrace{(\nabla U)_{\Delta V} \cdot d\mathbf{r}}_{-\mathbf{F}_e}$$

and so,

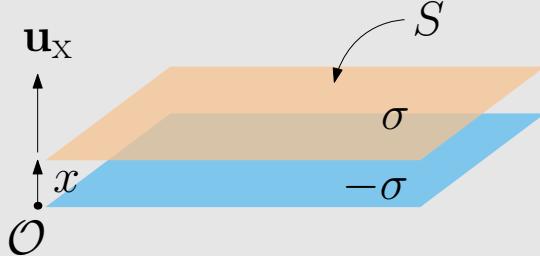
$$\mathbf{F} = -\Delta V \nabla Q(\mathbf{r}) + (\nabla U)_{\Delta V}$$

but $(\nabla U)_{\Delta V} = \frac{1}{2}\Delta V \nabla Q(\mathbf{r})$, so finally

$$\mathbf{F} = -\frac{1}{2}\Delta V \nabla Q(\mathbf{r}) = -(\nabla U)_{\Delta V}$$

Example 3.3 - Force on a parallel plane capacitor

Consider a parallel plane capacitor in which the plates, of area S , are charged with a charge density $\pm\sigma$. What is the electrostatic force acting on the top plate when both plates are separated by a distance x ?



Solution

According to Eq. (3.14), an electrostatic pressure $P_e = \frac{\sigma^2}{2\epsilon_0}$ acts on the top plate so that the total electrostatic force acting on it is

$$\mathbf{F}_e = -P_e S \mathbf{u}_x = -\frac{\sigma^2 S}{2\epsilon_0} \mathbf{u}_x.$$

On the other hand, if the top plate is fixed at distance x from the bottom plate, an external force must oppose to this electrostatic force. Since the system is electrically isolated, this force is given by

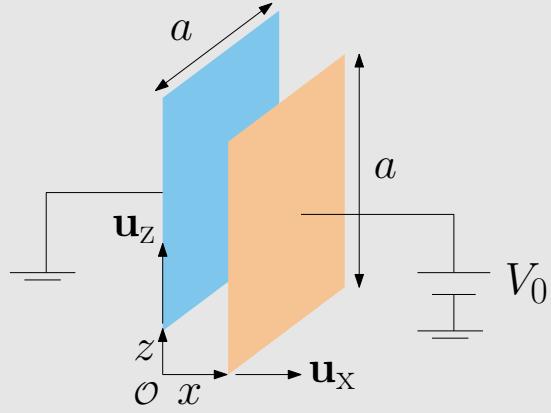
$$\mathbf{F} = (\nabla U)_Q = -\frac{Q^2}{2C(x)^2} \nabla C(\mathbf{r}),$$

where $Q = \sigma S$ is the charge on the top capacitor and $C(x) = \frac{\epsilon_0 S}{x}$ the capacitance. We find

$$\begin{aligned} \mathbf{F} &= -\frac{\sigma^2 S^2 x^2}{2(\epsilon_0 S)^2} \nabla \left(\frac{\epsilon_0 S}{x} \right) = -\frac{\sigma^2 S x^2}{2\epsilon_0} \left(-\frac{1}{x^2} \mathbf{u}_x \right) \\ &= \frac{\sigma^2 S}{2\epsilon_0} \mathbf{u}_x = -\mathbf{F}_e. \end{aligned}$$

Example 3.4 - Force between parallel planes at fixed potential

Two square, conductive and parallel plates of lateral size a are kept at a potential difference V_0 . The left plate is free to move in the $x - z$ plane. Neglecting edge effects, calculate the force needed to hold the left plate, and compare with the electrostatic force acting on the plate.

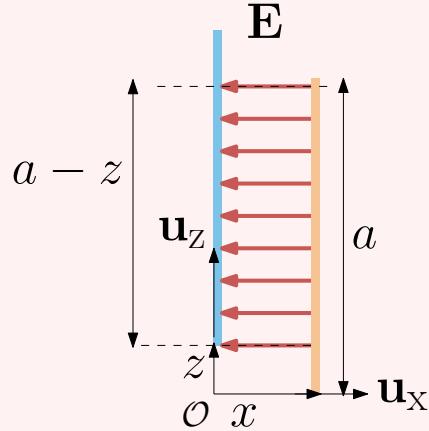


Solution

When the plates are at distance x and the bottom left corner of the left plate is at position $z\mathbf{u}_z$, the electrostatic energy of the system writes

$$U(x, z) = \frac{1}{2}C(x, z)V_0^2.$$

Neglecting edge effects, we assume that the electric field is homogeneous and that the field lines are perpendicular to both plates, confined to the region between the plates of area $a(a - z)$.



The capacity is then

$$C(x, z) = \frac{\epsilon_0 a(a - z)}{x}.$$

The force necessary to maintain the left plate at this position is then given by

$$\begin{aligned} \mathbf{F} &= -(\nabla U)_{\Delta V} = -\frac{\partial U}{\partial x}\mathbf{u}_x - \frac{\partial U}{\partial z}\mathbf{u}_z \\ &= \frac{V_0^2}{2} \left(\frac{\epsilon_0(a - z)a}{x^2}\mathbf{u}_x + \frac{\epsilon_0 a}{x}\mathbf{u}_z \right). \end{aligned}$$

Note that the electric field between the plates is given by

$$\mathbf{E} = -\frac{V_0}{x}\mathbf{u}_x = -\frac{\sigma}{\epsilon_0}\mathbf{u}_x$$

with $\sigma = \frac{\epsilon_0 V_0}{x}$ (resp. $-\sigma$) the charge density in the right (resp. left) plate. The electrostatic pressure then writes

$$P_e = \frac{\sigma^2}{2\epsilon_0} = \frac{\epsilon V_0^2}{2x^2}$$

and the electrostatic force on the left plate is

$$\mathbf{F}_e = a(a-z)P_e\mathbf{u}_x = \frac{V_0^2}{2} \frac{\epsilon_0(a-z)a}{x^2} \mathbf{u}_x.$$

3.7 Thermodynamic potentials for the forces acting on a conductor

In this section, we will see that it is possible to interpret the expressions for the force acting on a mobile conductor as deriving from thermodynamic potentials. This interpretation covers both cases of an isolated system and a system of conductors connected to batteries so that their potential is constant.

3.7.1 Free energy of a capacitor

The free energy F of a capacitor is defined as the constitution energy for the set of charges Q_1 , Q_2 , on each conductor. From chapter 2, we know that its numerical value is given by

$$F = \frac{1}{2}(Q_1 V_1 + Q_2 V_2)$$

and that it corresponds to the amount of external work (performed by an external operator), to bring the charges Q_1 , Q_2 from a reference state where they lie at zero potential to the final state with Q_1 on conductor 1 at potential V_1 and Q_2 on conductor 2 at potential V_2 . The basic assumption is that the operator exerts a force on each charge element dQ_1 , dQ_2 , that opposes at any time the electrostatic interaction force due to the charges already in their final state. In other words, the operator must work slowly enough so that the system can be assumed to be constituted in a quasi-static way. In thermodynamics, this corresponds to the situation where there is no heat dissipation and the system stores all the work of the operator. For the present

case, the operator's work is stored in the form of electrostatic energy. Both the definition and the terminology of *free energy* comes from thermodynamics: The external work between the reference state and the constituted (final) state is equal to the variation of the system free energy

The free energy is a *state function*: it does only depend on the state of the system. It does not depend on the way the system was assembled by the external operator. Mathematically, this means that F admits a total differential dF that represents the external work to bring the external charges dQ_1 and dQ_2 from their reference state (at zero potential) to their final state.

$$dF = \delta W \quad \text{external work to bring } dQ_1 \text{ and } dQ_2$$

Since the work of the operator reads $\delta W = V_1 dQ_1 + V_2 dQ_2$, we readily obtain an expression for dF

$$dF = V_1 dQ_1 + V_2 dQ_2.$$

Now, we wish to write an expression for the free energy state function $F(Q_1, Q_2)$. We thus need to integrate dF . In this aim, potentials are expressed as function of charges, using the inverse of the capacitance matrix $[\Gamma] = [C]^{-1}$ with coefficients Γ_{ij} :

$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad \Rightarrow \quad \Rightarrow \quad \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \quad (3.15)$$

From the condition that $dF = V_1 dQ_1 + V_2 dQ_2$ is a total differential, we know that it satisfies the Schwarz property

$$\left. \frac{\partial V_1}{\partial Q_2} \right|_{Q_1} = \left. \frac{\partial V_2}{\partial Q_1} \right|_{Q_2},$$

where the vertical bar and subscript $|_{Q_1}$ (resp. $|_{Q_2}$) indicates differentiation assuming constant Q_1 (resp. Q_2), leading to

$$\Gamma_{12} = \Gamma_{21},$$

that is, the matrix $[\Gamma]$ is symmetric¹⁰. We can thus express dF as a function of charges only

$$\begin{aligned} dF &= (\Gamma_{11}Q_1 + \Gamma_{12}Q_2)dQ_1 + (\Gamma_{21}Q_1 + \Gamma_{22}Q_2)dQ_2 \\ &= \Gamma_{11}Q_1dQ_1 + \Gamma_{12}(Q_2dQ_1 + Q_1dQ_2) + \Gamma_{22}Q_2dQ_2, \end{aligned}$$

and by integration, we find an expression for the free energy state function

$$F(Q_1, Q_2) = \frac{1}{2} \left(\Gamma_{11}Q_1^2 + 2\Gamma_{12}Q_1Q_2 + \Gamma_{22}Q_2^2 \right).$$

Finally, we recognize in the above equation, after factorizing Q_1 and Q_2 , the expressions for the potentials and we can write the free energy as

$$F = \frac{1}{2}(Q_1V_1 + Q_2V_2).$$

¹⁰Consequently, the capacitance matrix $[C]$ is symmetric too because the inverse of a symmetric matrix is symmetric. We retrieve the result $C_{12} = C_{21}$ from energy considerations.

We retrieved the expression for constitution energy of the system of conductors.

While the above expression gives the correct numerical value for the free energy, it should be kept in mind that potentials and charges are conjugated variables linked by the equation of state (3.15), in the same way as pressure and volume are conjugated variables in thermodynamics. As will be seen in the next section, care must be taken when F has to be differentiated, since V_1 and V_2 vary with Q_1 and Q_2 and vice-versa. Hence, a wisdom rule is to first express F as a function of its natural variables Q_1 and Q_2 .

As an example, consider the plane capacitor with $Q_2 = -Q_1$ (We define $Q_1 \equiv Q$ and $V_1 - V_2 \equiv \Delta V$). The free energy reads

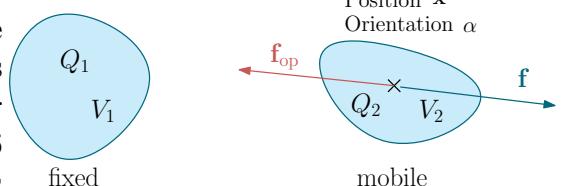
$$F = \frac{1}{2}(Q_1V_1 + Q_2V_2) = \frac{1}{2}Q_1(V_1 - V_2) = \frac{1}{2}Q\Delta V$$

Introducing the capacitance, we see that the free energy corresponds to the familiar expression for the energy stored in a capacitor:

$$Q = C\Delta V \quad \Rightarrow \quad F(Q) = \frac{1}{2} \frac{Q^2}{C}.$$

3.7.2 Forces on a conductor by means of the free energy

In this section, we consider an isolated system of conductors. Conductor 1 is fixed. It carries a charge Q_1 and is at potential V_1 . Conductor 2 carries a charge Q_2 and is at potential V_2 . It is mobile and can thus move in all directions around its center of mass \mathbf{r} or rotate in three dimensions. Conductor 2 thus has 6 freedom degrees, corresponding to three coordinates for its position and three angles α for its orientation.



The charges on the conductors exerts forces on the mobile conductor. An operator keeps the mobile conductor at rest and thus exerts a force \mathbf{f}_{op} and a torque $\boldsymbol{\tau}_{op}$ that are opposite to the force \mathbf{f} and torque $\boldsymbol{\tau}$ exerted by the charges. The free energy state function for the system is a function of the charges and the position and orientation variables of the mobile conductor $F(Q_1, Q_2, x, \alpha)$. We assume a single position variable x and a single orientation variable α only to simplify notations.

The free energy of the system corresponds to the external work used on the one hand to move charge elements dQ_1 and dQ_2 from a reference state at zero potential to their final state at V_1 and V_2 , respectively, and on the other hand to the work of the operator to maintain the mobile conductor at rest: $\delta W = \delta W_g + \delta W_{op}$. The free energy variation the reads

$$\begin{aligned} dF &= V_1dQ_1 + V_2dQ_2 + f_{x,op}dx + \tau_{\alpha,op}d\alpha, \\ &= V_1dQ_1 + V_2dQ_2 - f_xdx - \tau_\alpha d\alpha, \end{aligned}$$

where f_x and τ_α are the electrostatic force and torque components on the mobile conductor, along the x -axis and around the α -rotation axis. The fact that the free energy is a state function then translates into the mathematical property:

$$f_x = - \left. \frac{\partial F}{\partial x} \right|_{Q_i, \alpha}, \quad \tau_\alpha = - \left. \frac{\partial F}{\partial \alpha} \right|_{Q_i, x},$$

which correspond to expressions for the force and torque acting on the mobile conductor. In summary, to find the electrostatic force or torque acting on a mobile isolated conductor, the free energy of the system of conductors must be first expressed as a function of the charges on each conductor, the position and orientation coordinates of the mobile conductor. The force component along a given direction is the derivative of the free energy with respect to the corresponding coordinate, holding constant all other variables. The torque around a given axis is the derivative of the free energy with respect to the corresponding angular coordinate, holding constant all other variables.

For instance, for the mobile plate (along the x -axis) of a plane capacitor, the free energy reads

$$F(Q, x) = \frac{1}{2} \frac{Q^2}{C(x)}.$$

The force acting on the mobile plate then writes

$$f_x = \frac{1}{2} \frac{Q^2}{C^2} \frac{dC}{dx}.$$

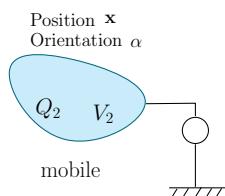
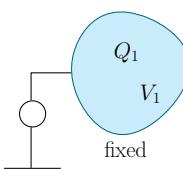
3.7.3 Another thermodynamic potential: the free enthalpy

In this section, another state function, the *free enthalpy*, is introduced and shown to be useful to calculate forces or torques acting on a mobile conductor when the system is no longer isolated but connected to generators holding the potentials constant.

We have seen that the differential for the free energy reads

$$dF = V_1 dQ_1 + V_2 dQ_2 - f_x dx - \tau_\alpha d\alpha,$$

allowing us to remember that (i) Q_i and V_i are conjugated variables and (ii) F is a natural function of charges (and freedom degrees of the mobile conductor, x and α).



The *free enthalpy* G is defined from the free energy F by a common operation in thermodynamics and classical mechanics, called a *Legendre transformation*. It amounts to subtracting

the product of a pair of conjugated variable: $F(Q_1, Q_2, x, \alpha) \rightarrow G(V_1, V_2, x, \alpha)$

Free enthalpy:

$$G = F - Q_1V_1 - Q_2V_2$$

It is a matter of calculus to obtain the differential of the free enthalpy:

$$\begin{aligned} dG &= dF - Q_1dV_1 - V_1dQ_1 - Q_2dV_2 - V_2dQ_2 \\ &= -Q_1dV_1 - Q_2dV_2 - f_xdx - \tau_\alpha d\alpha. \end{aligned}$$

From the above expression, we obtain the force and torque applied to the mobile conductor, assuming we have an integrated expression for the free enthalpy expressed as a function of its natural variables $G(V_1, V_2, x, \alpha)$:

$$f_x = -\left.\frac{\partial G}{\partial x}\right|_{V_i, \alpha}, \quad \tau_\alpha = -\left.\frac{\partial G}{\partial \alpha}\right|_{V_i, x}.$$

For instance, for a system of two conductors, the expressions for the free energy and the Legendre transform leading to the free enthalpy yield an expression for G :

$$F = \frac{1}{2}(Q_1V_1 + Q_2V_2) \quad \rightarrow \quad G = -\frac{1}{2}(Q_1V_1 + Q_2V_2).$$

Then, expanding the equation of state (3.15) by using the coefficients of the capacitance matrix, G can be expressed as a function of the potential difference

$$G = -\frac{1}{2}(C_{11}V_1^2 + C_{22}V_2^2 + 2C_{12}V_1V_2).$$

If the system of conductors is a capacitor, $Q_1 = -Q_2 \equiv Q$ and $V_1 - V_2 = \Delta V$, which yields the simplified expression

$$G = -\frac{1}{2}C(\Delta V)^2.$$

From this expression, we can derive the force on the mobile conductor:

$$f_x = -\left.\frac{\partial G}{\partial x}\right|_{V_i} = \frac{1}{2}\frac{\partial C}{\partial x}(\Delta V)^2.$$

The final result is the same as that obtained using the free energy, as can be seen by using the relation $(\Delta V) = Q/C$. However, care must be taken in each case to express the free energy as a function of charges or the free enthalpy as a function of potentials.

3.8 Summary and essential formulas

- In a metal in the presence of an electric field (out of static equilibrium), electrons will acquire a non-zero average velocity \mathbf{v} , which according to the Drude model is given by

$$\mathbf{v} = \frac{-e\tau}{m_e} \mathbf{E} = \mu \mathbf{E},$$

where $-e$ is the electron charge, τ the mean time between two collisions, and m_e the electron mass. This allows us to define the mobility $\mu = -\frac{e\tau}{m_e} < 0$.

- For a conductor Ω in electrostatic equilibrium, the electric field and the charge density vanish at every point inside Ω : $\rho = 0$, $\mathbf{E} = 0$. In the region outside the conductor, at a point infinitely close to the surface $\partial\Omega$, the electric field is given by Coulomb's theorem

$$\mathbf{E}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\epsilon_0} \mathbf{n} \quad \mathbf{x} \in \partial\Omega$$

where σ is the charge density on the surface of the conductor and \mathbf{n} the normal direction pointing towards the exterior of Ω .

- The electric potential is therefore constant at any point inside a conductor in static equilibrium, and equal to its value at its surface V_0 . The electrostatic energy of a conductor with total charge Q writes

$$U = \frac{1}{2} V_0 Q.$$

- In a system of N conductors in electrostatic equilibrium, the charge Q_i at the surface of the conductor i depends linearly on the potentials of the N conductors

$$Q_i = \sum_{j=1}^N C_{ij} V_j,$$

defining the capacitance matrix $[C_{ij}]$. For the case $N = 2$,

$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \underbrace{\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}}_{\text{Capacitance matrix}} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}.$$

If in addition, $Q_1 = -Q_2 = Q$, the two conductors constitute a capacitor, of capacitance

$$C = \frac{Q}{\Delta V},$$

where ΔV is the potential difference between the conductors.

- The electrostatic energy associated to a capacitor is given by

$$U = \frac{1}{2}Q\Delta V = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}C(\Delta V)^2$$

- The capacity for a plane capacitor of surface S and distance e between the plates is

$$C = \frac{\epsilon_0 S}{e}$$

Unit: Farad (F).

- The electrostatic force \mathbf{F}_e and torque $\boldsymbol{\tau}(O)$ acting on a conductor Ω at equilibrium are given by integrals over its surface $\partial\Omega$ of the electrostatic pressure P_e :

$$P_e = \frac{\sigma^2}{2\epsilon_0}, \quad \mathbf{F}_e = \iint_{\partial\Omega} P_e(\mathbf{r})\mathbf{n}(\mathbf{r})dS(\mathbf{r}) \quad \text{and} \quad \boldsymbol{\tau}(O) = \iint_{\partial\Omega} \mathbf{r} \times p_e(\mathbf{r})\mathbf{n}(\mathbf{r})dS(\mathbf{r}).$$

- The free energy $F(Q_1, Q_2, x, \alpha)$ for an isolated system of two conductors corresponds to the stored electrostatic energy. The free energy can be evaluated as

$$F(Q_1, Q_2, x, \alpha) = \frac{1}{2}(Q_1V_1 + Q_2V_2)$$

where Q_i and V_i are conjugated variables satisfying $V_i = \sum_j \Gamma_{ij}Q_j$, with Γ_{ij} , the coefficients of the inverse of the capacitance matrix.

- The free enthalpy for a system of two conductors held at constant potentials reads

$$G(V_1, V_2, x, \alpha) = F - Q_1V_1 - Q_2V_2 = -\frac{1}{2}(Q_1V_1 + Q_2V_2).$$

- The differentials of the free energy and free enthalpy for a system of two conductors read

$$\begin{aligned} dF &= V_1dQ_1 + V_2dQ_2 - f_xdx - \tau_\alpha d\alpha, \\ dG &= -Q_1dV_1 - Q_2dV_2 - f_xdx - \tau_\alpha d\alpha \end{aligned}$$

and the force and torque applied on a mobile conductor are given by

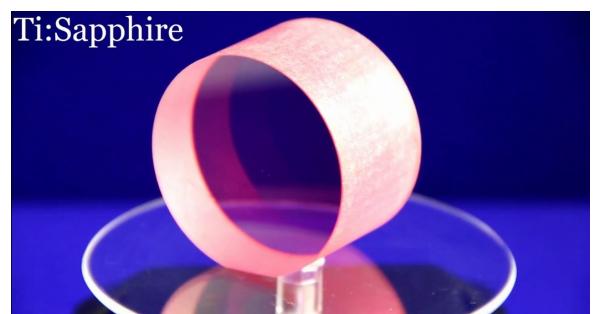
$$f_x = -\left.\frac{\partial F}{\partial x}\right|_{Q_i} = -\left.\frac{\partial G}{\partial x}\right|_{V_i}, \quad \tau_\alpha = -\left.\frac{\partial F}{\partial \alpha}\right|_{Q_i} = -\left.\frac{\partial G}{\partial \alpha}\right|_{V_i}.$$

- Free energy and free enthalpy for a capacitor:

$$F = \frac{1}{2}\frac{Q^2}{C}, \quad G = -\frac{1}{2}C\Delta V^2.$$

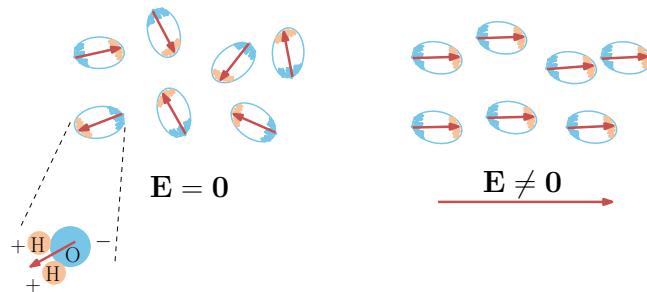
Chapter 4

The field in dielectric media



Introduction

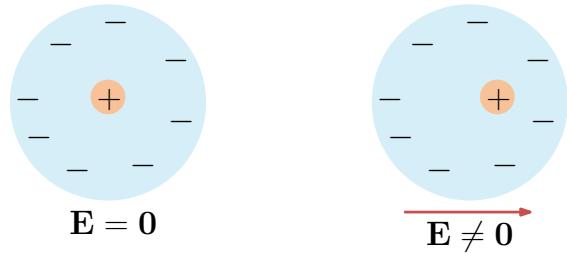
In this chapter we will discuss the behavior of insulating (also known as dielectric) materials. Within an insulator, electrons are strongly bound to the nuclei and cannot move freely. Since atoms and molecules tend to be electrically neutral, one might think that insulators do not interact significantly with an external electrostatic field. However, this is not true. Michael Faraday discovered that by filling the space between two conductors with a dielectric, the capacity of the system increases. We will show below that, indeed, molecules and atoms in insulating materials may produce an electric field that tends to weaken the total electric field seen by a free charge inside the dielectric.



As we have seen in chapter 2, there are polar molecules (for example water, shown in the figure above) for which negative and positive charges are sufficiently displaced so that they are

assimilated to an electric dipole. In the case of water, the electrons are, in average, concentrated near the oxygen atom, leaving an excess of positive charge near the hydrogen atoms, and so a water molecule has a permanent dipole moment. Normally, in the absence of an electric field, each molecule has a random orientation resulting in a zero net dipole moment. In contrast, when an external electric field is present, a torque is exerted on the molecules and they align so that their dipole moments point in the direction of the external field, producing therefore a net dipole moment in the material. This means that molecules in matter will produce a significant, macroscopic electric field in response to the applied field.

The same phenomenon can occur in an insulator composed of non-polar molecules. Under the effect of an electric field, the negative and positive charge densities of each molecule will be shifted with respect to each other due to the electrostatic force, so that a dipole moment appears in each molecule as a consequence of the applied electric field.



These molecules, once polarized, are also aligned with their dipole moments pointing in the direction of the electric field. In any case, once an external electric field is applied to an insulator, a net dipole moment appears at the macroscopic level. It is then necessary to consider its contribution to the total electric field that will be present.

4.1 Polarization vector

If a medium is polarized, that is, if locally the molecules have their dipole moments pointing all in the same direction, every infinitesimal element of volume is then characterized by an infinitesimal net electric dipole moment $d\mathbf{p}$.

Definition 4.1: Polarization

The polarization $\mathbf{P}(\mathbf{x})$ of a medium at a point \mathbf{x} is defined as the average electrical dipole moment density at point \mathbf{x} . In other words, in an infinitesimal volume d^3x around \mathbf{x} , one has

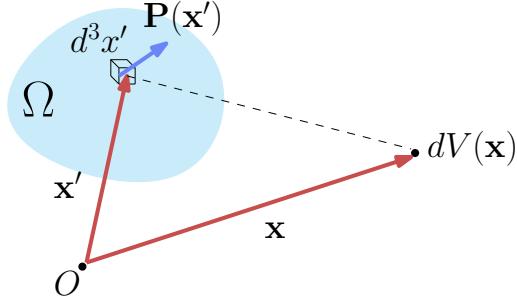
$$\mathbf{P}(\mathbf{x})d^3x = d\mathbf{p}(\mathbf{x}),$$

where $d\mathbf{p}(\mathbf{x})$ corresponds to the total dipole moment contained in d^3x . The polarization vector has units of Cm^{-2} .

We consider then a polarized material as a continuous medium so that \mathbf{P} is defined at every point by a smooth function of \mathbf{x} . For this, one considers around any point \mathbf{x} a volume element d^3x that is very small from the macroscopic point of view, yet big enough so that it contains a large number of molecules. A full justification of this averaging process will be discussed in section 4.8.

4.1.1 Potential generated by a dielectric medium

Consider a volume Ω of dielectric characterized by its polarization $\mathbf{P}(\mathbf{x}')$ at every point \mathbf{x}' of Ω .



Each volume element d^3x' behaves like a macroscopic dipole moment $d\mathbf{p}(\mathbf{x}') = \mathbf{P}(\mathbf{x}')d^3x'$. The potential at \mathbf{x} due to this volume element will then be the potential of an electric dipole,

$$dV(\mathbf{x}) = \frac{d\mathbf{p}(\mathbf{x}') \cdot (\mathbf{x} - \mathbf{x}')}{4\pi\epsilon_0|\mathbf{x} - \mathbf{x}'|^3} = \frac{\mathbf{P}(\mathbf{x}') \cdot (\mathbf{x} - \mathbf{x}')d^3x'}{4\pi\epsilon_0|\mathbf{x} - \mathbf{x}'|^3}.$$

The total potential at \mathbf{x} is obtained by integration over the volume Ω

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{\mathbf{P}(\mathbf{x}') \cdot (\mathbf{x} - \mathbf{x}')d^3x'}{|\mathbf{x} - \mathbf{x}'|^3}. \quad (4.1)$$

4.1.2 Bound charge densities

Now we show that the potential given by Eq. (4.1) can be written as the potential generated by a charged volume Ω , that is, as an integral of the general form that yields the potential for a charge distribution with an adequate charge distribution. Using the identity

$$\nabla \frac{1}{|\mathbf{x}|} = -\frac{\mathbf{x}}{|\mathbf{x}|^3},$$

we can write

$$\frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = \nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|},$$

where ∇' denotes the gradient with respect to coordinates of \mathbf{x}' . The potential can then be rewritten as

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \mathbf{P}(\mathbf{x}') \cdot \nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|} d^3x'.$$

Now consider the identity: $\nabla \cdot (f\mathbf{F}) = f\nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla f$, where f is a scalar field and \mathbf{F} a vector field. Taking $f = 1/|\mathbf{x} - \mathbf{x}'|$ and $\mathbf{F} = \mathbf{P}(\mathbf{x}')$, we obtain

$$\mathbf{P}(\mathbf{x}') \cdot \nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|} = \nabla' \cdot \frac{\mathbf{P}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} - \frac{1}{|\mathbf{x} - \mathbf{x}'|} \nabla' \cdot \mathbf{P}(\mathbf{x}').$$

With this,

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \nabla' \cdot \frac{\mathbf{P}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' - \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{1}{|\mathbf{x} - \mathbf{x}'|} \nabla' \cdot \mathbf{P}(\mathbf{x}') d^3x'.$$

The first integral can be transformed into an integral over the closed surface $\partial\Omega$ by means of Green-Ostrogradsky's theorem, leading to

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \oint_{\partial\Omega} \frac{\mathbf{P}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \cdot d\mathbf{S}(\mathbf{x}') - \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{\nabla' \cdot \mathbf{P}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

Or, equivalently,

$$V(\mathbf{x}) = \oint_{\partial\Omega} \frac{\mathbf{P}(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}')}{4\pi\epsilon_0 |\mathbf{x} - \mathbf{x}'|} + \iiint_{\Omega} \frac{-\nabla' \cdot \mathbf{P}(\mathbf{x}') d^3x'}{4\pi\epsilon_0 |\mathbf{x} - \mathbf{x}'|}. \quad (4.2)$$

We see that the potential generated by a dielectric material, given by Eq. (4.2), is generated by a charge distribution with volume density $\rho_P(\mathbf{x}) = -\nabla \cdot \mathbf{P}(\mathbf{x})$ (C m^{-3}) in Ω and surface density $\sigma_P(\mathbf{x}) = \mathbf{P}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$ (C m^{-2}) in $\partial\Omega$. These quantities are called bound (or polarization) charge densities. This terminology refers to the fact that these charges are tantamount to globally neutral pairs of negative and positive charges, the negative charge being close to the positive charge, where *close* means at a much smaller distance compared to the distance from the charges to the point at which the elementary dipole is generating the potential. For instance in a crystal, electrons are bound to the positive ions that constitute the lattice of the crystal.

Definition 4.2: Bound charge density

A polarized dielectric Ω can be seen as a medium containing a bound charge (to distinguish it from the free charge density) with volume density ρ_P in Ω and surface density σ_P in $\partial\Omega$. The volume density of bound charges ρ_P is a measure of the non-uniformity of \mathbf{P} within the material

$$\rho_P(\mathbf{x}) = -\nabla \cdot \mathbf{P}(\mathbf{x}) \quad \mathbf{x} \in \Omega, \quad (4.3)$$

while the surface density of bound charges is given by the component of \mathbf{P} that is normal to the surface,

$$\sigma_P(\mathbf{x}) = \mathbf{P}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega. \quad (4.4)$$

Remark

Note that the total bound charge of a dielectric is

$$Q_P = \iiint_{\Omega} -\nabla' \cdot \mathbf{P}(\mathbf{x}') d^3x' + \iint_{\partial\Omega} \mathbf{P}(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}') = 0.$$

It is shown to be null due to the divergence theorem and in agreement with the fact that a dielectric is globally neutral. The appearance of a dipole moment at the molecular level comes from a charge distribution that is only locally different from zero and it is this bound charge that is taken into account in ρ_P and σ_P .

In summary, the potential due to the dielectric material is

$$V(\mathbf{x}) = \iint_{\partial\Omega} \frac{\sigma_P(\mathbf{x}') dS(\mathbf{x}')}{4\pi\epsilon_0 |\mathbf{x} - \mathbf{x}'|} + \iiint_{\Omega} \frac{\rho_P(\mathbf{x}') d^3x'}{4\pi\epsilon_0 |\mathbf{x} - \mathbf{x}'|}. \quad (4.5)$$

The electric field is obtained by using the identity $\mathbf{E} = -\nabla V$. Naturally, one obtains a Coulomb integral with the appropriate charge distributions:

$$\mathbf{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iint_{\partial\Omega} \frac{\sigma_P(\mathbf{x}') (\mathbf{x} - \mathbf{x}') dS(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} + \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{\rho_P(\mathbf{x}') (\mathbf{x} - \mathbf{x}') d^3x'}{|\mathbf{x} - \mathbf{x}'|^3}.$$

4.2 Gauss's law in a dielectric

Suppose that we now have a certain distribution of *free* charges ρ_{free} submerged in a dielectric medium, that is, charges that are not already counted in polarization charges¹. These charges generate an electric field that will polarize the molecules in the medium surrounding them.

In addition to these free charges, a polarized dielectric medium will also contribute to the total electric field via its bound charge density ρ_P . Gauss's law remains valid in a dielectric, but one must include all the charge density on the right hand side (free and bound charge) to write

$$\nabla \cdot \mathbf{E}(\mathbf{x}) = \frac{\rho_{\text{free}}(\mathbf{x}) + \rho_P(\mathbf{x})}{\epsilon_0}. \quad (4.6)$$

The bound charge density can be written in terms of the polarization vector as $\rho_P = -\nabla \cdot \mathbf{P}$ according to Eq. (4.3), so that

$$\nabla \cdot \mathbf{E}(\mathbf{x}) = \frac{\rho_{\text{free}}(\mathbf{x}) - \nabla \cdot \mathbf{P}(\mathbf{x})}{\epsilon_0}$$

and rearranging the terms, we obtain

$$\nabla \cdot (\epsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x})) = \rho_{\text{free}}.$$

We see that the field $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ is such that its divergence has only the free charges as sources.

¹the term *free* can be misleading since ρ_{free} may not only contain charges that are free to move but also immobile charges such as ions or charged impurities that are not free to move but are not bound to any other charge of opposite sign

Definition 4.3: Electric displacement and generalized Gauss's law in a dielectric

The electric displacement is the vector field defined as

$$\mathbf{D}(\mathbf{x}) = \epsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x}), \quad (4.7)$$

which satisfies the differential Gauss law

$$\nabla \cdot \mathbf{D}(\mathbf{x}) = \rho_{\text{free}}(\mathbf{x}), \quad (4.8)$$

whose integral form is

$$\iint_S \mathbf{D}(\mathbf{x}') \cdot d\mathbf{S}(\mathbf{x}') = Q_{\text{free}}(S), \quad (4.9)$$

so that the flux of the field \mathbf{D} through any closed surface S equals the free charge enclosed by it.

The physical unit for the displacement vector is the same as for the polarization, that is, $[\text{C m}^{-2}]$. It represents a surface charge density.

Remarks

1. Equation (4.8) is a more convenient way of writing Eq. (4.6). Note that the displacement field \mathbf{D} has the advantage of conserving only free charges as sources, making it easier to deal with Eq. (4.8) instead of Eq. (4.6).
2. All the information about the polarization of the medium is contained in \mathbf{D} . In vacuum, we have $\mathbf{P} = \mathbf{0}$, $\mathbf{D} = \epsilon_0 \mathbf{E}$ and Eq. (4.9) is equivalent to Gauss's law in vacuum.
3. In reality, there is no fundamental distinction between free and bound charges at the atomic scale. Gauss's law in a dielectric, Eq. (4.8), and the distinction between ρ_{free} and ρ_P comes in reality after averaging both the charge density and the resulting electric field so as to smooth their rapid variations at the atomic scale. This will be detailed in section 4.8.

4.3 Electric susceptibility and dielectric constant

Consider a dielectric medium where the distribution of free charges $\rho_{\text{free}}(\mathbf{x})$ is known at any point \mathbf{x} . If we would like to find the electric field \mathbf{E} and displacement field \mathbf{D} at any point in

the dielectric, we have to find the solution to Maxwell's equations

$$\nabla \times \mathbf{E}(\mathbf{x}) = \mathbf{0}, \quad (4.10)$$

$$\nabla \cdot \mathbf{D}(\mathbf{x}) = \rho_{\text{free}}(\mathbf{x}). \quad (4.11)$$

The first equation tells us that the electric field is conservative, $\mathbf{E} = -\nabla V$. If we were looking for a solution in vacuum, $\mathbf{D} \equiv \epsilon_0 \mathbf{E}$ and $\rho(\mathbf{x}) = \rho_{\text{free}}(\mathbf{x})$ as there is no polarization charge, hence the second equation would become $\nabla \cdot \epsilon_0 \mathbf{E}(\mathbf{x}) = \rho_{\text{free}}(\mathbf{x})$, leading to Poisson's equation

$$\Delta V = -\rho/\epsilon_0.$$

Thus, the scalar potential can be obtained by solving Poisson's equation, and the electric field could be inferred from the potential. The situation is different in a dielectric since the system of equations (4.10),(4.11) is not closed. A relation between \mathbf{E} and \mathbf{D} is required in order to apply a similar procedure as in vacuum and derive the electric field, the displacement field, and the polarization from the knowledge of free charges only. This relation plays the role of a constitutive equation for the medium, or an equation of state, telling us how the dielectric responds to the electric field, that is, what is the polarization of the medium in response to an applied electric field. This is a medium dependent relation and we will now specify certain classes of equation of states $\mathbf{D}(\mathbf{E})$ (or equivalently $\mathbf{P}(\mathbf{E})$).

4.3.1 Linear homogeneous and isotropic media

For a wide variety of materials, the properties are *homogeneous*, that is, are uniform throughout the dielectric volume and do not depend on \mathbf{x} .

For a wide variety of materials, the polarization density \mathbf{P} depends linearly on the electric field. The displacement vector $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ then also depends linearly on \mathbf{E} . This means that in a Cartesian basis $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$, a relation exists between the components of \mathbf{D} and \mathbf{E} :

$$\begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \underbrace{\begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix}}_{\text{permittivity matrix}} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

where the matrix $[\epsilon]$ is called the permittivity matrix. It has nine coefficients which are assumed to be independent of \mathbf{x} for an homogeneous medium, while the electric and displacement field can a priori depend on \mathbf{x} . We will show later than the permittivity matrix is in fact a symmetric matrix, leading to only six independant coefficients as $\epsilon_{xy} = \epsilon_{yx}$, $\epsilon_{xz} = \epsilon_{zx}$ and $\epsilon_{yz} = \epsilon_{zy}$. Furthermore, any symmetric matrix can be diagonalized: a specific vector basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ exists in which the permittivity matrix is diagonal:

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

The axes corresponding to a diagonal permittivity matrix are the principal axes of the dielectric crystal. The medium is anisotropic if $\epsilon_1 \neq \epsilon_2 \neq \epsilon_3$. This is the case for instance in birefringent crystals such as calcite, where rays are refracted in different directions



Now for an isotropic medium (for which all directions are equivalent), all the diagonal coefficients of the permittivity are equal $\epsilon_1 = \epsilon_2 = \epsilon_3 \equiv \epsilon$. This means that the linear relation between \mathbf{D} and \mathbf{E} takes the simple form

$$\mathbf{D}(\mathbf{x}) = \epsilon \mathbf{E}(\mathbf{x}),$$

where the permittivity ϵ is now a scalar quantity that is independent of \mathbf{x} for a homogeneous dielectric.

4.3.2 Electric susceptibility and dielectric constant

For a linear and isotropic medium (relaxing the assumption of uniformity), \mathbf{P} is thus written in the form

$$\mathbf{P}(\mathbf{x}) = \epsilon_0 \chi(\mathbf{x}) \mathbf{E}(\mathbf{x}),$$

where the dimensionless quantity χ is called the *electric susceptibility* of the material. Thus, the expression of \mathbf{D} in an isotropic medium is, according to Eq. (4.7),

$$\mathbf{D}(\mathbf{x}) = \epsilon_0(1 + \chi(\mathbf{x})) \mathbf{E}(\mathbf{x}) = \epsilon(\mathbf{x}) \mathbf{E}(\mathbf{x}). \quad (4.12)$$

Definition 4.4: Dielectric constant.

We define $\epsilon = \epsilon_0(1 + \chi)$ as the permittivity of the medium, which for an isotropic and homogeneous material (χ independent of position) is simply a constant. The relative permittivity ϵ_r , also called dielectric constant, is defined as the dimensionless quantity

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} \quad (4.13)$$

so that

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \epsilon_r \mathbf{E}.$$

Remarks

1. In terms of the permittivity, the polarization vector can be written as

$$\mathbf{P}(\mathbf{x}) = (\epsilon - \epsilon_0)\mathbf{E}(\mathbf{x})$$

From the latter expression one sees that, at any point inside a dielectric with $\rho_{\text{free}} = 0$, the volume density of bound charge $\rho_P = -\nabla \cdot \mathbf{P} = -\nabla \cdot [(\epsilon - \epsilon_0)\mathbf{E}] = -\frac{\epsilon_0}{\epsilon}(\nabla\epsilon) \cdot \mathbf{E}$ is zero unless there is a gradient of permittivity in the direction of the electric field². In an homogeneous dielectric, $\rho_P = 0$ and the bound charge is thus constrained to appear on its surface.

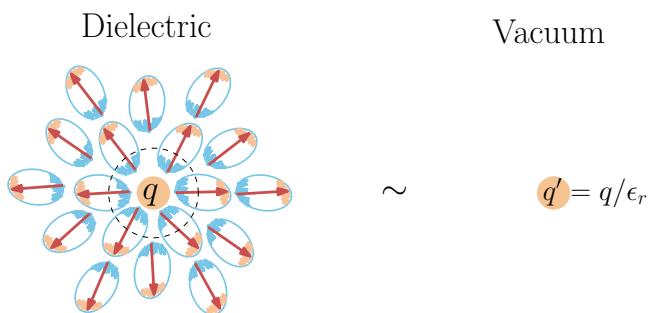
2. The dielectric constant ϵ_r is always greater than or equal to 1 (vacuum). In the presence of a free charge density ρ_{free} and in an homogeneous, isotropic and linear material, the differential form of generalized Gauss's law can be rewritten in terms of the dielectric constant and the electric field as

$$\nabla \cdot \mathbf{E}(\mathbf{x}) = \frac{\rho_{\text{free}}(\mathbf{x})}{\epsilon} = \frac{\rho_{\text{free}}(\mathbf{x})}{\epsilon_r \epsilon_0}. \quad (4.14)$$

This means that in a linear, homogeneous and isotropic dielectric medium, the laws of electrostatics are equivalent to those in vacuum, provided that the permittivity is replaced by $\epsilon = \epsilon_r \epsilon_0$ or, equivalently, if the free charge density is replaced by $\rho_{\text{free}}/\epsilon_r$.

4.3.3 Physical interpretation of the dielectric constant

When a medium is polarized under an external electric field, it produces in turn an electric field that opposes the external one, so that the total electric field is weakened by a factor ϵ_r with respect to the value it would have in vacuum. This occurs because the bound charge density tends to screen the free charges. The following figure illustrates the case of a positive charge q in a dielectric medium. At the interface between the charge and the medium, a cloud of negative charge surrounds the charge q , producing a screening effect. The electric field generated in the medium is equivalent to the field of a charge $q/\epsilon_r < q$ in vacuum.



²We have used the fact that there is no free charge and Gauss's law states $\nabla \cdot \mathbf{E} = \rho_P/\epsilon_0$

A dielectric with $\epsilon_r \gg 1$ then behaves like a conductor, in the sense that the electric field generated by the bound charge tends to cancel the external field.

The following table gives examples of dielectric constants.

Material	Relative permittivity (dielectric constant ϵ_r)
Vacuum	1.0000
Air	1.0006
PTFE, FEP ("Teflon") ...	2.0
Polypropylene	2.20 to 2.28
ABS resin	2.4 to 3.2
Polystyrene	2.45 to 4.0
Waxed paper	2.5
Transformer oil	2.5 to 4
Hard Rubber	2.5 to 4.80
Wood (Oak)	3.3
Silicones	3.4 to 4.3
Bakelite	3.5 to 6.0
Quartz, fused	3.8
Wood (Maple)	4.4
Glass	4.9 to 7.5
Castor oil	5.0
Wood (Birch)	5.2
Mica, muscovite	5.0 to 8.7
Glass-bonded mica	6.3 to 9.3
Porcelain, Steatite	6.5
Alumina	8.0 to 10.0
Distilled water	80.0
Barium-strontium-titanite	7500

- The relative permittivity ϵ_r is a dimensionless constant
- $\epsilon_r \geq 1$
- $\epsilon_r \sim 1$ for gases
- Metals / conductors: $\epsilon_r \gg 1$
- In spite of its large relative permittivity, water is not a good dielectric for enhancing the capacity of a capacitor (mica can be used for this purpose)

Example 4.1 - Charged sphere in a dielectric

A sphere of radius R with an homogeneously distributed free charge q is submerged in a homogeneous dielectric medium of constant ϵ_r .

- Calculate the electric field and polarization vectors at a distance r from the sphere.
- Calculate the polarization charge density.

Solution

- Since the free charge distribution has a clear spherical symmetry and it is surrounded by a dielectric medium, the bound charge inherits the same symmetry. As a consequence, the symmetries are the same for \mathbf{D} and \mathbf{E} for a constant ϵ_r , and we conclude that $\mathbf{D} = D(r)\mathbf{u}_r$ in spherical coordinates (with the origin at the center of the sphere). Using Gauss's law for dielectrics, Eq. (4.9), with a spherical surface

S of radius r ($r > R$) and concentric to the sphere of charge q , we write

$$\iint_S \mathbf{D}(\mathbf{x}') \cdot d\mathbf{S}(\mathbf{x}') = 4\pi r^2 D(r) = q,$$

from which we obtain

$$\mathbf{D}(r) = \frac{q}{4\pi r^2} \mathbf{u}_r,$$

The electric field and polarization can be easily evaluated from \mathbf{D} :

$$\mathbf{D}(r) = \epsilon_r \epsilon_0 \mathbf{E}(r) \rightarrow \mathbf{E}(r) = \frac{q}{4\pi \epsilon_r \epsilon_0 r^2} \mathbf{u}_r.$$

Furthermore,

$$\mathbf{D}(\mathbf{x}) = \epsilon_0 \mathbf{E}(\mathbf{x}) + \mathbf{P}(\mathbf{x}) \quad \text{and} \quad \mathbf{D} = \epsilon \mathbf{E}.$$

Then,

$$\begin{aligned} \mathbf{P}(r) &= (\epsilon - \epsilon_0) \mathbf{E}(r) = \frac{q}{4\pi r^2} \mathbf{u}_r - \frac{\epsilon_0 q}{4\pi \epsilon_r \epsilon_0 r^2} \mathbf{u}_r \\ &= \frac{q}{4\pi r^2} \left(1 - \frac{1}{\epsilon_r}\right) \mathbf{u}_r = \frac{q(\epsilon_r - 1)}{4\pi \epsilon_r r^2} \mathbf{u}_r. \end{aligned}$$

- b.** Clearly, the dielectric has attenuated the magnitude of the electric field by a factor ϵ_r compared to its value in vacuum. The electric field is generated by all charges, that is, bound and free charges. The free charge is simply the charge of the sphere q , whereas the bound charge exists only at the interface between the dielectric and the spherical charge. Indeed, at all points of the dielectric

$$\nabla \cdot \mathbf{P}(\mathbf{x}) = (\epsilon - \epsilon_0) \underbrace{\nabla \cdot \mathbf{E}(\mathbf{x})}_{0} = 0$$

and so,

$$\rho_P(\mathbf{x}) = -\nabla \cdot \mathbf{P}(\mathbf{x}) = 0.$$

Using $\mathbf{P}(r) = q(\epsilon_r - 1)/(4\pi \epsilon_r r^2) \mathbf{u}_r$, the surface density of bound charge at $r = R$ will be

$$\sigma_P = \mathbf{P} \cdot \mathbf{n}|_{r=R} = \mathbf{P}(R) \cdot (-\mathbf{u}_r) = -\frac{(\epsilon_r - 1)q}{4\pi \epsilon_r R^2}$$

and the total bound charge at $r = R$ is

$$Q_S = -4\pi R^2 \frac{(\epsilon_r - 1)q}{4\pi \epsilon_r R^2} = -\frac{(\epsilon_r - 1)}{\epsilon_r} q.$$

This charge partially shields the free charge q . Note finally that the dielectric is globally neutral, so that a bound charge $-Q_S$ appears at the outer boundary of the dielectric, but has no influence on the electric field since this boundary is located at infinity.

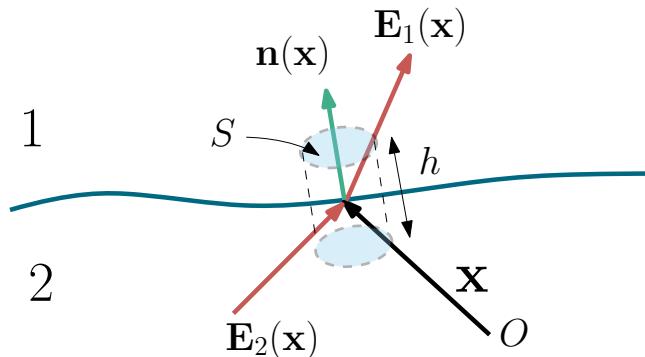
4.4 Boundary conditions at the interface between dielectric media

The conditions that the electric field must satisfy at an interface that separates two dielectric media are deduced from the differential laws

$$\begin{aligned}\nabla \cdot \mathbf{D}(\mathbf{x}) &= \rho(\mathbf{x}), \\ \nabla \times \mathbf{E}(\mathbf{x}) &= 0.\end{aligned}$$

4.4.1 The normal component of the displacement field is discontinuous at a charged interface

Consider two regions of space and a point \mathbf{x} at the interface between the two regions. A closed cylinder can be constructed as shown in the figure below, such that the normal to the cap in region 2 (resp. region 1) is oriented along $\mathbf{n}(\mathbf{x})$ (resp. $-\mathbf{n}(\mathbf{x})$) where $\mathbf{n}(\mathbf{x})$ is the normal to the interface at \mathbf{x} , pointing from medium 2 to medium 1.



Gauss's law (4.9) reads

$$\iint_S \mathbf{D}(\mathbf{x}) \cdot d\mathbf{S}(\mathbf{x}) = \iiint_{\Omega(S)} \rho(\mathbf{x}) d^3x.$$

If h is the height of the cylinder and R the radius of its circular cap, the lateral area of the cylinder is $2\pi hR$. In the limit when $h \rightarrow 0$, the flux of \mathbf{D} through the side tends to zero, and only the flux through the caps must be considered, the surfaces of which are equal to $A = \pi R^2$:

$$\lim_{h \rightarrow 0} \iint_S \mathbf{D}(\mathbf{x}') \cdot d\mathbf{S}(\mathbf{x}') = \mathbf{n}(\mathbf{x}) \cdot (\mathbf{D}_1(\mathbf{x}) - \mathbf{D}_2(\mathbf{x})) A.$$

On the other hand, if we admit that the interface carries a surface density of free charges $\sigma(\mathbf{x})$, then the free charge enclosed in the Gauss surface when $h \rightarrow 0$ is simply $A\sigma(\mathbf{x})$. We then obtain

$$\mathbf{n}(\mathbf{x}) \cdot (\mathbf{D}_1(\mathbf{x}) - \mathbf{D}_2(\mathbf{x})) A = \sigma(\mathbf{x}) A.$$

Thus, at every point \mathbf{x} at the interface between two regions, the normal component of the field $\mathbf{D}(\mathbf{x})$ is discontinuous if the interface is charged. The discontinuity satisfies

$$\mathbf{n} \cdot (\mathbf{D}_1 - \mathbf{D}_2)|_{\mathbf{x}} = \sigma(\mathbf{x}), \quad (4.15)$$

where \mathbf{n} is the normal vector that points from medium 2 to medium 1. For linear, homogeneous and isotropic dielectric media, this is rewritten in terms of the electric field as

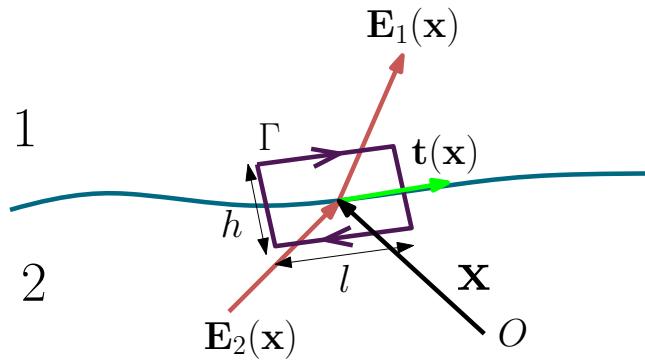
$$\mathbf{n} \cdot (\epsilon_1 \mathbf{E}_1 - \epsilon_2 \mathbf{E}_2)|_{\mathbf{x}} = \sigma(\mathbf{x}). \quad (4.16)$$

4.4.2 The tangential component of the electric field is continuous at an interface

Now we will see that the tangential component of the electric field is continuous at the interface between two regions. For this, we can integrate the equation

$$\nabla \times \mathbf{E}(\mathbf{x}) = 0,$$

on a closed rectangular path Γ , as shown in the figure, resulting in the law of circulation for the electrostatic field $\oint_{\Gamma} \mathbf{E}(\mathbf{x}) \cdot d\mathbf{x} = 0$.



In the case where $h \rightarrow 0$, the circulation is reduced to that over the segments parallel to the tangential direction $\mathbf{t}(\mathbf{x})$ to the interface at \mathbf{x} ,

$$l(\mathbf{E}_1(\mathbf{x}) - \mathbf{E}_2(\mathbf{x})) \cdot \mathbf{t}(\mathbf{x}) = 0.$$

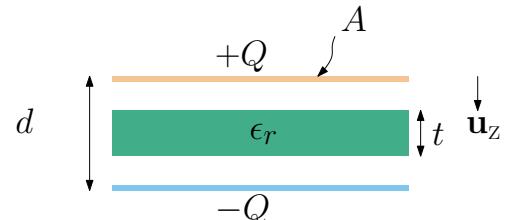
It follows that the tangential component of $\mathbf{E}(\mathbf{x})$ is continuous at any point \mathbf{x} of the interface between two dielectric media,

$$\mathbf{t} \cdot (\mathbf{E}_1 - \mathbf{E}_2)|_{\mathbf{x}} = 0. \quad (4.17)$$

4.5 Capacitors with dielectrics

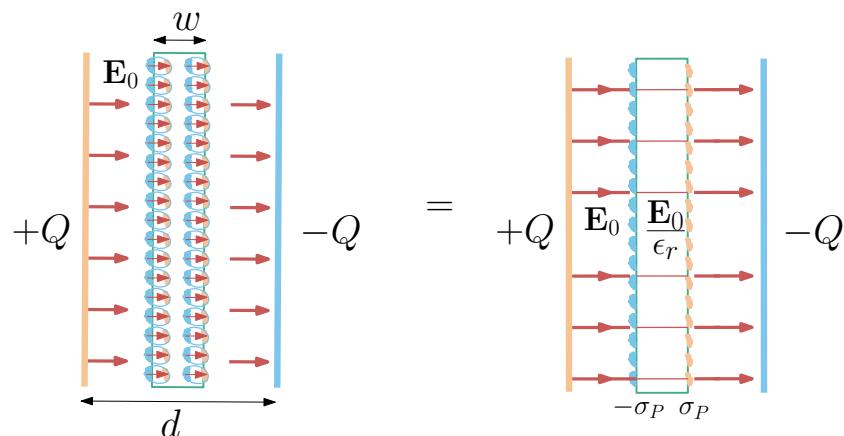
4.5.1 Capacity of a capacitor with dielectric medium

Now we will see that when a dielectric material is inserted between the two plates of a capacitor, the capacitance increases by a factor ϵ_r , where ϵ_r is the dielectric constant, as observed by the physicist Michael Faraday in the 19th century.



Consider the case of a flat capacitor, of area A , separation distance d between the plates, and with a charge Q . A dielectric material of dielectric constant ϵ_r and of width $w \leq d$ is placed between the plates. To find the capacitance, let us first calculate the potential difference between the plates. In the absence of dielectric, we know that the magnitude of the electric field between the plates (assuming that the lateral extension of the plates is much larger than the thickness d) is given by $E_0 = \frac{\sigma}{\epsilon_0} = \frac{Q}{\epsilon_0 A}$. Gauss's law in a dielectric, Eq. (4.9), tells us that the free charges are sources of a displacement field \mathbf{D} , so that the problem is equivalent to solve for the electric field in vacuum after multiplying the charges by ϵ_0 . The \mathbf{D} field generated by the plates is then homogeneous and equal to $\mathbf{D} = \sigma \mathbf{u}_x$ in the region between the plates, and null outside. At any point outside the dielectric, $\epsilon_r = 1$ (vacuum), hence we have $\mathbf{D} = \epsilon_0 \mathbf{E}$ and the electric field will be, as in the case without dielectric, $\mathbf{E}_0 = (\sigma/\epsilon_0) \mathbf{u}_x$.

Now, inside the dielectric, $\mathbf{D} = \epsilon_r \epsilon_0 \mathbf{E}$, and then the electric field is weakened by a factor ϵ_r with respect to its value in vacuum, $\mathbf{E}_D = \mathbf{E}_0 / \epsilon_r$. Since \mathbf{E}_D is a uniform field inside the dielectric, the polarization charges only appear on the surface of the latter. Physically, the orientation of the electric dipoles along the electric field generated by the capacitor is at the origin of a polarization charge at the boundaries of the dielectric that shields the field of the plates of the capacitor at all points inside the dielectric, as shown in the following figure.



The potential can be found by integrating the electric field on a vertical curve (parallel to \mathbf{u}_x) that joins the upper plate to the lower one

$$\Delta V = \int_0^d E(x)dx = E_0(d-w) + E_D w = \frac{Q}{A\epsilon_0}(d-w) + \frac{Q}{A\epsilon_0\epsilon_r}w = \frac{Q}{A\epsilon_0} \left(d - w \left(1 - \frac{1}{\epsilon_r} \right) \right).$$

From this equation, we obtain the capacitance

$$C = \frac{Q}{\Delta V} = \frac{\epsilon_0 A}{d - w \left(1 - \frac{1}{\epsilon_r} \right)}.$$

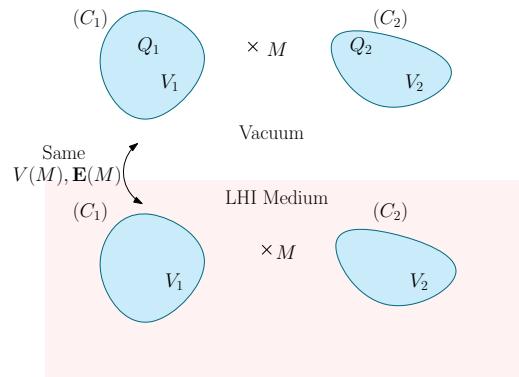
Note that when $w = d$ (i.e., the space between plates is filled with dielectric),

$$C = \frac{\epsilon_r \epsilon_0 A}{d} = \epsilon_r C_0,$$

where C_0 is the capacity in the absence of dielectric. Inserting a dielectric of large dielectric constant between the plates of a capacitor is a standard way to increase the capacity of the capacitor without changing its geometry.

4.5.2 Capacitance coefficients in a dielectric medium

Consider the more general case of a system of N conductors in a linear homogeneous and isotropic medium, exemplified with $N = 2$ in the figure on the right. We established the relation between charges Q_i carried by the conductors at potential V_i when they are in vacuum. This led to the definition of the capacitance matrix. Does the capacitance matrix change when the conductors are in a dielectric medium rather than in vacuum?



We know from previous section that for a linear homogeneous and isotropic dielectric inserted in a capacitor, the capacity is the product of the capacity in vacuum by the relative permittivity of the medium. We expect a similar change in capacitance coefficients for a system of N conductors. The figure compares the situation with $N = 2$ conductors in vacuum at potential V_1 and V_2 with that of the same conductors at the same positions and potentials but a dielectric medium replaced vacuum.

We know that between the conductors in vacuum, the potential at any point M must satisfy Laplace's equation

$$\Delta V = 0,$$

together with boundary conditions

$$V(M) = V_1 \text{ for } M \in (C_1), \text{ and } V(M) = V_2 \text{ for } M \in (C_2).$$

The field is then obtained from the potential:

$$\mathbf{E} = -\nabla V.$$

If a linear homogeneous and isotropic dielectric medium is present between the conductors, the displacement field between the conductors satisfies the Maxwell-Gauss equation with no free charge between the conductors:

$$\nabla \cdot \mathbf{D} = 0.$$

Since $\mathbf{D} = \epsilon \mathbf{E}$, the Maxwell-Gauss equation leads to $\epsilon \nabla \cdot \mathbf{E} = 0$ and using the fact that the electric field is conservative, $\mathbf{E} = -\nabla V$, we find that the potential satisfies Poisson's equation in this case too, $\Delta V = 0$. The boundary conditions are exactly the same as in vacuum, therefore, the potential at any point between the conductors is the same as in vacuum.

A difference between the two cases arises for the free charge carried by the conductors. In vacuum, the Coulomb theorem states that at the surface of a conductor, $\epsilon_0 \mathbf{E} = \sigma_0 \mathbf{n}$, where \mathbf{E} denotes the field at any point of a conductor, \mathbf{n} the normal to the conductor and σ_0 the free charge density on the surface of the conductor. In the dielectric, the Coulomb theorem is equivalent to applying the jump condition for the normal component of the displacement field, the tangential component being zero due to the continuity of the tangential component of the electric field at the interface between a conductor and the dielectric. Hence,

$$\mathbf{D} - \mathbf{0} = \epsilon \mathbf{E} = \sigma \mathbf{n}$$

Since the electric field and the normal are the same as in vacuum, we find

$$\sigma = \epsilon \frac{\sigma_0}{\epsilon_0} = \epsilon_r \sigma_0,$$

that is, the free charge density on the conductors is multiplied by ϵ_r . In conclusion, the free charges on the conductors takes their value in vaccum multiplied by ϵ_r and since the potentials are the same, the relation that defines the capacitance matrix

$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

shows that in the presence of a dielectric medium between the conductors,

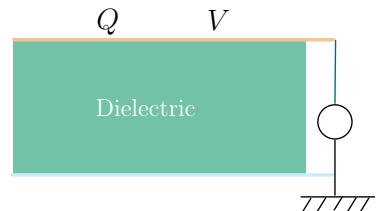
the coefficients of the capacitance matrix are multiplied by ϵ_r .

4.6 The electrostatic free energy for a polarized medium

In chapter 2, we defined and calculated the electrostatic potential energy of a charge distribution as the work necessary for an external operator to bring the charges of the distribution from infinity, or ground, where the potential is zero, to their actual position in a quasistatic way. We have identified this work to the free energy F of the system, which is a state function depending only on the final state of the system. The question we consider here is how this free energy is modified for polarization charges?

4.6.1 Electrostatic free energy density for a dielectric inserted in a plane capacitor

Consider a dielectric that is under the action of an external electric field, or under the action of external (free) electric charges. For instance, we can think about the dielectric between the plates of a capacitor carrying charge Q on its plate at potential V , the other plate being grounded.



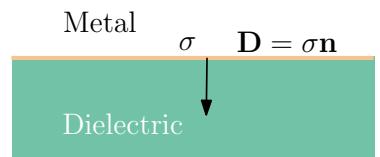
The free energy of the system corresponds by definition to the work of the external operator, that is the generator, which supplies charges to polarize the dielectric medium, this work being made in a quasistatic way, without spilling energy into heat. We know that F is a state function, and we have shown in chapter 2 that in order to supply a charge dQ to the positive plate of the capacitor starting from an established equilibrium state characterized by potential V and charge Q , therefore reaching the neighbor state $V + dV$, $Q + dQ$, the work of the generator, and so the infinitesimal variation of the free energy reads

$$dF = V dQ.$$

The potential V and the charge Q are conjugated variables in the above expression, hence, as shown in chapter 2, the variation of the free energy must be expressed as a function of Q only if we wish to integrate the free energy $F(Q)$. However, we are here interested in the dependence of the free energy on polarization charges and we will therefore express dF as a function of the polarization \mathbf{P} via the electric and displacement fields, \mathbf{E} and \mathbf{D} .

The displacement field is easily obtained from the jump relation of its normal component at the interface between the dielectric and the metal plate, where $\mathbf{D} = \mathbf{0}$:

$$\mathbf{D} = \sigma \mathbf{n},$$



where \mathbf{n} denotes the normal to the metal plate pointing toward the dielectric. By differentiation, this relation yields

$$dD = d\sigma = \frac{dQ}{S}$$

Since $V = Ee$, where e denotes the thickness dielectric (capacitor), we can express the variation in the free energy as a product of the volume of the dielectric in the capacitor by the variation of the free energy density:

$$dF = VdQ = \underbrace{Ve}_{V} \underbrace{\frac{Sd\sigma}{dQ}}_{\text{volume}} = \underbrace{eS}_{\text{volume}} \underbrace{E \frac{d\sigma}{dD}}_{\text{volume}} = \underbrace{(eS)}_{\text{volume}} \underbrace{(EdD)}_{df}.$$

We finally identify the free energy density, whose differential reads

$$df = \mathbf{E} \cdot d\mathbf{D} .$$

The free energy density is an exact differential, showing that \mathbf{E} and \mathbf{D} are conjugated variables. This means that the natural variable for the free energy density is \mathbf{D} . As exemplified below, the differential df can then be expressed as a function of \mathbf{D} only before integration if we wish to find an expression for the state function $f(\mathbf{D})$.

4.6.2 Free energy density for a dielectric medium - general case

The result obtained in the particular case of the capacitor, where the electric and displacement fields are uniform between the plates can be generalized to the case of any polarized dielectric medium in the presence of an electric field $\mathbf{E}(\mathbf{x})$ by simply considering $\mathbf{E}(\mathbf{x})$ and $\mathbf{D}(\mathbf{x})$ as locally uniform over an infinitesimal volume d^3x .

$$df = \mathbf{E}(\mathbf{x}) \cdot d\mathbf{D}(\mathbf{x}).$$

Hence, an integration of the variation of free energy density over the volume Ω of the dielectric yields the free energy variation

$$dF = \iiint_{\Omega} \mathbf{E}(\mathbf{x}) \cdot d\mathbf{D}(\mathbf{x}) d^3x ,$$

which is an exact differential, F being a state function. The presence of two infinitesimal elements $d\mathbf{D}$ and d^3x in the above expression might seem weird. However, it is perfectly correct as the volume integral will eliminate d^3x and result in a single infinitesimal element on each side of this equation. From the definition of the displacement vector, $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, we find by differentiation

$$d\mathbf{D} = \epsilon_0 d\mathbf{E} + d\mathbf{P} \Rightarrow df = \underbrace{\epsilon_0 \mathbf{E} \cdot d\mathbf{E}}_{\substack{\text{energy to} \\ \text{establish } \mathbf{E}}} + \underbrace{\mathbf{E} \cdot d\mathbf{P}}_{\substack{\text{energy to} \\ \text{establish } \mathbf{P}}}$$

The first term would be present even without dielectric medium. It is recognized as the electrostatic energy to establish the electric field, that is, to change the electric field from \mathbf{E} to $\mathbf{E} + d\mathbf{E}$. The second term corresponds to the energy to establish the polarization in the medium, that is, to change the polarization from \mathbf{P} to $\mathbf{P} + d\mathbf{P}$ while maintaining the electric field \mathbf{E} .

4.6.3 Free energy density for a linear homogeneous and isotropic medium

How can we find the free energy density of a polarized dielectric once we know its differential? Remember that \mathbf{E} and \mathbf{P} are not independent. They are linked by an equation of state. For instance in a linear homogeneous and isotropic medium, $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$. Hence

$$df = \epsilon_0 \mathbf{E} \cdot d\mathbf{E} + \epsilon_0 \chi \mathbf{E} \cdot d\mathbf{E}$$

A dielectric with an established electric field \mathbf{E} therefore carries the electrostatic free energy density

$$f = \epsilon_0 \frac{\mathbf{E}^2}{2} + \epsilon_0 \chi \frac{\mathbf{E}^2}{2},$$

and an integration over the volume of the dielectric yields the free energy of the system. Again the first term on the right-hand side corresponds to the energy to establish the electric field, as if only free charges were present in vacuum, and the second term to the energy to establish the polarization. Adding both contributions leads to the equivalent forms for the free energy, for a linear homogeneous and isotropic dielectric:

$$F = \iiint_{\Omega} \epsilon \frac{\mathbf{E}^2}{2} d^3x = \iiint_{\Omega} \frac{\mathbf{D}^2}{2\epsilon} d^3x = \iiint_{\Omega} \frac{\mathbf{E} \cdot \mathbf{D}}{2} d^3x = \iiint_{\Omega} \epsilon_0 \frac{\mathbf{E}^2}{2} d^3x + \iiint_{\Omega} \frac{\mathbf{E} \cdot \mathbf{P}}{2} d^3x.$$

These relations are equivalently demonstrated by starting from the reference state with zero electric and displacement fields, $\mathbf{E} = \mathbf{0}$ and $\mathbf{D} = \mathbf{0}$, and integrating the free energy up to the final state characterized by \mathbf{E} and \mathbf{D} , with $\mathbf{D} = \epsilon \mathbf{E}$ for a linear homogeneous and isotropic medium. Integration is performed along the path of actual equilibrium states characterized by the electric field $\alpha \mathbf{E}$ and displacement $\alpha \mathbf{D}$, where $0 \leq \alpha \leq 1$. The variation of the free energy to reach a neighbor state $((\alpha + d\alpha) \mathbf{E}, (\alpha + d\alpha) \mathbf{D})$ from an actual state $(\alpha \mathbf{E}, \alpha \mathbf{D})$ is $df = \alpha \mathbf{E} \cdot (d\alpha) \mathbf{D} = \mathbf{E} \cdot \mathbf{D} d\alpha$. Hence when α varies³ between 0 and 1,

$$f = \int_0^1 \mathbf{E} \cdot \mathbf{D} d\alpha = \mathbf{E} \cdot \mathbf{D} \int_0^1 \alpha d\alpha \frac{1}{2} \mathbf{E} \cdot \mathbf{D}.$$

The density of free energy for a linear homogeneous and isotropic medium writes

$$f = \frac{1}{2} \mathbf{E} \cdot \mathbf{D} = \epsilon \frac{\mathbf{E}^2}{2} = \epsilon_0 \frac{\mathbf{E}^2}{2} + \frac{\mathbf{E} \cdot \mathbf{P}}{2}.$$

³Note that \mathbf{E} and \mathbf{D} are notations characterizing the final state of the dielectric. They do not change along the integration path.

If we express it as a function of its natural variable \mathbf{D} , it reads,

$$f(\mathbf{D}) = \frac{1}{2\epsilon} \mathbf{D}^2.$$

4.6.4 Free enthalpy

Another thermodynamic state function is often useful: the density of free enthalpy, also called the Gibbs free energy, obtained from the density of free energy via a Legendre transform:

$$g = f - \mathbf{E} \cdot \mathbf{D}.$$

Differentiating this definition leads to the differential for the volume density of free enthalpy

$$dg = -\mathbf{D} \cdot d\mathbf{E},$$

which shows that the natural variable for the free enthalpy $g(\mathbf{E})$ is the electric field \mathbf{E} . Using the Legendre transform together with expression for the free energy as a function of \mathbf{E} for a linear homogeneous and isotropic dielectric leads to

$$g(\mathbf{E}) = -\frac{1}{2}\epsilon\mathbf{E}^2$$

To illustrate the use of the free enthalpy, consider again a linear and possibly anisotropic dielectric with permittivity matrix $[\epsilon]$. From the equation of state expressed in a Cartesian basis,

$$\begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix},$$

the differential for the Gibbs free energy reads

$$dg = -\mathbf{D} \cdot d\mathbf{E} = -D_x dE_x - D_y dE_y - D_z dE_z$$

Now the fact that f is a state function means that df is an exact differential and therefore follows the Schwarz property

$$\left. \frac{\partial D_x}{\partial E_y} \right|_{E_x} = \left. \frac{\partial D_y}{\partial E_x} \right|_{E_y}$$

that is,

$$\epsilon_{xy} = \epsilon_{yx}.$$

Similarly, considering the two other pairs of variables (E_x, E_z) and (E_y, E_z) , we can show that $\epsilon_{xz} = \epsilon_{zx}$ and $\epsilon_{yz} = \epsilon_{zy}$. Thus, using the properties of the Gibbs free energy provides allows us to infer in an elegant way that the permittivity matrix is symmetric.

4.7 Forces applied to a dielectric medium

4.7.1 Force and torque density

We already know that the force \mathbf{F} exerted on a dipole moment \mathbf{p} in an electric field \mathbf{E} can be expressed component-wise, in a Cartesian basis, as

$$F_x = \mathbf{p} \cdot \nabla E_x, \quad F_y = \mathbf{p} \cdot \nabla E_y, \quad F_z = \mathbf{p} \cdot \nabla E_z$$

and the torque on the dipole moment reads

$$\mathbf{T} = \mathbf{p} \times \mathbf{E}.$$

For a volume element d^3x in a dielectric medium, the dipole moment is $d\mathbf{p}(\mathbf{x}) = \mathbf{P}(\mathbf{x})d^3x$, thus the infinitesimal force exerted on this dipole moment reads, component-wise, $dF_x = \mathbf{P}d\tau \cdot \nabla E_x$, leading to the force density $f_x = \frac{dF_x}{d^3x}$, that is,

$$f_x = \mathbf{P} \cdot \nabla E_x, \quad f_y = \mathbf{P} \cdot \nabla E_y, \quad f_z = \mathbf{P} \cdot \nabla E_z,$$

or in compact format,

$$\mathbf{f} = (\mathbf{P} \cdot \nabla) \mathbf{E}.$$

The torque density exerted on a dielectric medium is obtained similarly and writes

$$\boldsymbol{\tau} = \mathbf{P} \times \mathbf{E}.$$

These results call for a couple of remarks:

- the force components f_x, f_y, f_z are zero if \mathbf{E} is uniform: Forces apply in areas where the electric field \mathbf{E} is non-uniform.
- In a Linear homogeneous and isotropic medium, the torque density $\boldsymbol{\tau} = \mathbf{0}$ since $\mathbf{P} \parallel \mathbf{E}$.
- For a linear homogeneous and isotropic medium, $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$, and the force becomes

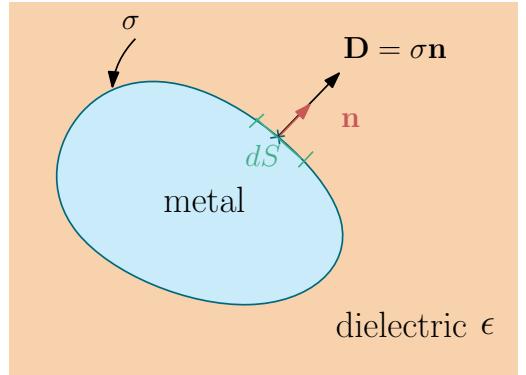
$$\mathbf{f} = \frac{\epsilon_0}{2} \nabla (\chi \mathbf{E}^2) = \frac{1}{2} \nabla (\mathbf{E} \cdot \mathbf{P}).$$

It is important to appreciate that this expression differs from that obtained for a permanent dipole moment, that does not depend on \mathbf{E} : $\mathbf{f} = \nabla(\mathbf{P} \cdot \mathbf{E})$.

4.7.2 Electrostatic pressure for a linear homogeneous isotropic medium

We extend here the concept of electrostatic pressure to the case of a conductor surrounded by a dielectric medium, supposed to be linear, homogeneous and isotropic, with a permittivity ϵ .

The conductor carries a charge density σ on its surface. In order to derive an expression for the electrostatic pressure, we consider an infinitesimal surface (dS) on the conductor and rest of the system (R). We are looking for the effect of (R) on (dS), that is, the force $d\mathbf{f}$ exerted on the charge $dq = \sigma dS$ on (dS) by electrostatic forces due to charges in (R).



$$d\mathbf{f} = dq \mathbf{E}_{(R)}(M) = \sigma dS \mathbf{E}_{(R)}(M),$$

where M is a point on the surface of the conductor and $\mathbf{E}_{(R)}(M)$ denotes the electric field at M generated by charges in (R), that is, all charges except those on (dS) which do not act on themselves.

Coulomb's theorem allows us to find the displacement field $\mathbf{D} = \sigma \mathbf{n}$, or equivalently the total electric field at M , produced by charges on (dS) and charges in (R): $\mathbf{E} = \frac{\sigma}{\epsilon} \mathbf{n}$. We can write the field at M as the superposition of the fields produced by charges on (dS) and by charges in (R):

$$\mathbf{E}(M) = \mathbf{E}_{(dS)}(M) + \mathbf{E}_{(R)}(M).$$

Now from the point of view of M , the infinitesimal surface (dS) behaves like an infinite plane carrying a charge of surface density σ , so the the field $\mathbf{E}_{(dS)}(M)$ is the same as the field produced by an infinite plane in the vicinity of the plane:

$$\mathbf{E}_{(dS)}(M) = \frac{\sigma}{2\epsilon} \mathbf{n}.$$

From these relations, we infer the field $\mathbf{E}_{(R)}(M)$

$$\mathbf{E}_{(R)}(M) = \mathbf{E}(M) - \mathbf{E}_{(dS)}(M) = \frac{\sigma}{\epsilon} \mathbf{n} - \frac{\sigma}{2\epsilon} \mathbf{n}, \quad \text{i.e.,} \quad \boxed{\mathbf{E}_{(R)}(M) = \frac{\sigma}{2\epsilon} \mathbf{n}.}$$

We can then express the force exerted on (dS) in the form of a product of the surface area by the electrostatic pressure:

$$d\mathbf{f}(M) = \underbrace{\sigma dS \frac{\sigma}{2\epsilon} \mathbf{n}}_{p_e(M) dS \mathbf{n}} \Rightarrow \boxed{p_e(M) = \frac{\sigma^2(M)}{2\epsilon}}.$$

We see that the electrostatic pressure takes the same form as in vacuum, but the permittivity is replaced by that of the medium. The electrostatic pressure is therefore reduced by a factor ϵ_r , in conjunction with the fact that inserting a dielectric medium in a capacitor allows us to increase its capacity by a factor ϵ_r .

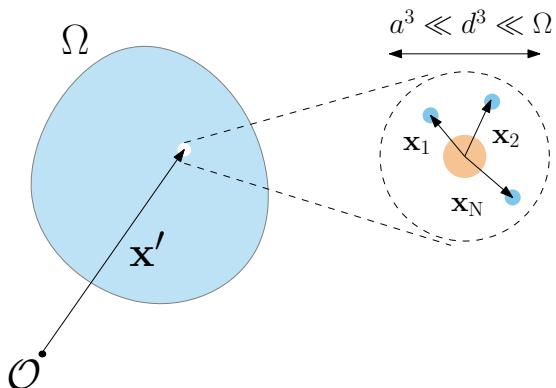
4.8 Gauss's law: from vacuum to a dielectric

In this chapter, we have treated an insulating material as a continuous volume where one can define a polarization density at any point. In reality, any charge density and the resulting electric field have singularities at the atomic scale and both are always related by Gauss's law in vacuum, $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$. We will show that the generalized Gauss law for a dielectric material, $\nabla \cdot \mathbf{D} = \rho_{\text{free}}$, can be obtained by performing a spatial average of both the charge density and the electric field.

Let us recall that the average of a quantity f is given by

$$\langle f(\mathbf{x}) \rangle = \iiint_{\mathbb{R}^3} f(\mathbf{x}') \varphi(\mathbf{x} - \mathbf{x}') d^3x' \quad (4.18)$$

where φ is a function centered at the origin with a spatial extension L and unit integral. In chapter 1, it was shown that by averaging the charge density of an atom of size a over a scale $d \gg a$, one can justify the use of the Coulomb integral for a charged volume $\Omega \gg d^3$. However, this approximation is somehow *crude* in the sense that φ is considered to be a constant within the atom. We will show now that by considering the first order variations of φ at the atomic scale, one can retrieve the behavior of dielectrics.



Let us consider an atom whose center of mass is located at \mathbf{x}' and constituted of N charges q_i located at \mathbf{x}_i relative to \mathbf{x}' . The average charge density of this atom is

$$\langle \rho_{\text{at}}(\mathbf{x}) \rangle = \sum_{i=1}^N q_i \varphi(\mathbf{x} - \mathbf{x}' - \mathbf{x}_i).$$

Instead of completely neglecting the spatial variation of φ at the scale of the atom (as in chapter 1), we will consider here the first order approximation

$$\varphi(\mathbf{x} - \mathbf{x}' - \mathbf{x}_i) \approx \varphi(\mathbf{x} - \mathbf{x}') - \nabla \varphi(\mathbf{x} - \mathbf{x}') \cdot \mathbf{x}_i,$$

so that

$$\langle \rho_{\text{at}} \rangle(\mathbf{x}) = \underbrace{\left(\sum_{i=1}^N q_i \right)}_{Q_{\text{at}}} \varphi(\mathbf{x} - \mathbf{x}') - \underbrace{\left(\sum_{i=1}^N q_i \mathbf{x}_i \right)}_{\mathbf{p}_{\text{at}}} \cdot \nabla \varphi(\mathbf{x} - \mathbf{x}') = Q_{\text{at}} \varphi(\mathbf{x} - \mathbf{x}') - \mathbf{p}_{\text{at}} \cdot \nabla \varphi(\mathbf{x} - \mathbf{x}'),$$

with Q_{at} and \mathbf{p}_{at} the total charge and dipole moment of the atom, respectively. While in general atoms are neutral, $Q_{\text{at}} \neq 0$ when dealing with ions or with charged materials for which there is an excess (or lack) of electrons per atom. Finally, by summing over all the atoms of the dielectric, one obtains the total average charge distribution

$$\langle \rho \rangle(\mathbf{x}) = \sum_k Q_k \varphi(\mathbf{x} - \mathbf{x}_k) - \sum_k \mathbf{p}_k \cdot \nabla \varphi(\mathbf{x} - \mathbf{x}_k).$$

Since \mathbf{p}_k is not a vector field, we can write

$$\mathbf{p}_k \cdot \nabla \varphi(\mathbf{x} - \mathbf{x}_k) = \nabla \cdot (\mathbf{p}_k \varphi(\mathbf{x} - \mathbf{x}_k)).$$

By defining the polarization vector \mathbf{P} (which is an averaged dipole moment density) as

$$\mathbf{P}(\mathbf{x}) = \sum_k \mathbf{p}_k \varphi(\mathbf{x} - \mathbf{x}_k)$$

and by identifying $\sum_k Q_k \varphi(\mathbf{x} - \mathbf{x}_k)$ as the free charge density $\langle \rho_{\text{free}} \rangle(\mathbf{x})$ (the name is justified when dealing with free electrons in metals, although in general it may also contain immobile charges within the material such as fixed ions or charged impurities), we obtain

$$\langle \rho \rangle(\mathbf{x}) = \langle \rho_{\text{free}} \rangle(\mathbf{x}) - \nabla \cdot \mathbf{P}(\mathbf{x}),$$

where we recognize the bound charge density $\rho_P = -\nabla \cdot \mathbf{P}$. By spatially averaging Gauss's law in vacuum

$$\langle \nabla \cdot \mathbf{E} \rangle(\mathbf{x}) = \frac{\langle \rho(\mathbf{x}) \rangle}{\epsilon_0} = \frac{\langle \rho_{\text{free}} \rangle(\mathbf{x}) - \nabla \cdot \mathbf{P}(\mathbf{x})}{\epsilon_0}$$

and since φ is zero at infinity, it is easy to show that,

$$\begin{aligned} \langle \nabla \cdot \mathbf{E}(\mathbf{x}) \rangle &= \iiint_{\mathbb{R}} [\nabla' \cdot \mathbf{E}(\mathbf{x}')] \varphi(\mathbf{x} - \mathbf{x}') d^3x' = - \iiint_{\mathbb{R}} \mathbf{E}(\mathbf{x}') \cdot \nabla' \varphi(\mathbf{x} - \mathbf{x}') d^3x' \\ &= \iiint_{\mathbb{R}} \mathbf{E}(\mathbf{x}') \cdot \nabla \varphi(\mathbf{x} - \mathbf{x}') d^3x' = \nabla \cdot \left(\iiint_{\mathbb{R}} \mathbf{E}(\mathbf{x}') \varphi(\mathbf{x} - \mathbf{x}') d^3x' \right) = \nabla \cdot \langle \mathbf{E} \rangle(\mathbf{x}). \end{aligned}$$

We have ended up demonstrating that the spatially averaged electric field satisfies

$$\nabla \cdot \langle \mathbf{E} \rangle = \frac{\langle \rho_{\text{free}} \rangle - \nabla \cdot \mathbf{P}}{\epsilon_0} \quad (4.19)$$

and that by defining the displacement vector as

$$\mathbf{D} = \epsilon_0 \langle \mathbf{E} \rangle + \mathbf{P},$$

one retrieves Gauss's law for dielectrics,

$$\nabla \cdot \mathbf{D} = \rho_{\text{free}}. \quad (4.20)$$

4.9 Summary and essential formulas

- A material medium, composed of molecules or neutral atoms, is characterized by its polarization vector \mathbf{P} , which represents the average dipole moment per unit volume.
- A polarized medium Ω for which $\mathbf{P} \neq \mathbf{0}$ can be seen as a charge distribution in Ω , consisting of a volume density ρ_P and a surface density σ_P on $\partial\Omega$, and given by

$$\rho_P = -\nabla \cdot \mathbf{P} \quad \text{and} \quad \sigma_P = \mathbf{P} \cdot \mathbf{n},$$

where \mathbf{n} is the normal to the surface $\partial\Omega$.

- The potential at any point \mathbf{x} generated by a polarized dielectric carrying both a distribution of free charges of density ρ_{free} and polarization charges of density ρ_P in the volume Ω and σ_P on its surface $\partial\Omega$ writes

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \left[\iiint_{(\Omega)} \frac{\rho_{\text{free}}(\mathbf{x}') + \rho_P(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \iint_{\partial\Omega} \frac{\sigma_P(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} dS' \right].$$

- Gauss's law in a material medium is written

$$\iint_{\partial\Omega} \mathbf{D} \cdot d\mathbf{S} = Q_{\text{free}} \quad \Leftrightarrow \quad \nabla \cdot \mathbf{D} = \rho_{\text{free}},$$

where \mathbf{D} is the displacement vector, defined by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}.$$

- The displacement field exhibits a discontinuity of its normal component across an interface

$$(\mathbf{D}_1 - \mathbf{D}_2) \cdot \mathbf{n} = \sigma,$$

where σ denotes the surface density of free charge at the interface and \mathbf{n} the normal to the interface pointing from medium 2 to medium 1.

- The tangential component of the electric field is continuous across an interface between two dielectric media

$$(\mathbf{E}_1 - \mathbf{E}_2) \cdot \mathbf{t} = 0$$

- The differential for the volume density of electrostatic free energy writes and for the electrostatic free energy read, respectively,

$$df = \mathbf{E} \cdot d\mathbf{D}, \quad \text{and} \quad dF = \iiint_{\Omega} \mathbf{E}(\mathbf{x}') \cdot d\mathbf{D}(\mathbf{x}') d^3x'.$$

Integration to obtain the free energy F requires an equation of state $\mathbf{D}(\mathbf{E})$.

- The force and torque exerted on a polarized medium in an electric field are expressed via the force and torque densities

$$\mathbf{f} = (\mathbf{P} \cdot \nabla) \mathbf{E} \quad \text{and} \quad \boldsymbol{\tau} = \mathbf{P} \times \mathbf{E}.$$

The torque is non-zero only if \mathbf{P} and \mathbf{E} are not colinear.

- In a linear and isotropic medium, the dipole moments are aligned according to the orientation of the electric field \mathbf{E} , and the polarization density \mathbf{P} is proportional to \mathbf{E} , $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$, with $\chi > 0$. The displacement vector can be written

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

where $\epsilon = \epsilon_0(1 + \chi) > \epsilon_0$ is the permittivity of the medium. If the medium is homogeneous, everything occurs as in vacuum, except that the permittivity is multiplied by the relative permittivity $\epsilon_r = 1 + \chi \geq 1$. In particular:

- The fundamental equations of electrostatics in matter are then written

$$\nabla \cdot \mathbf{E} = \frac{\rho_{\text{free}}}{\epsilon} \quad \text{and} \quad \nabla \times \mathbf{E} = \mathbf{0},$$

where ρ_{free} represents the density of free charges, the contribution of the bound (polarization) charges is thus included in ϵ . The dielectric medium tends to decrease the magnitude of the electric field generated by free charges, because the induced bound charges tend to shield the latter.

- The capacity of the plane capacitor with a dielectric of permittivity ϵ between the plates reads

$$C = \frac{\epsilon S}{e}.$$

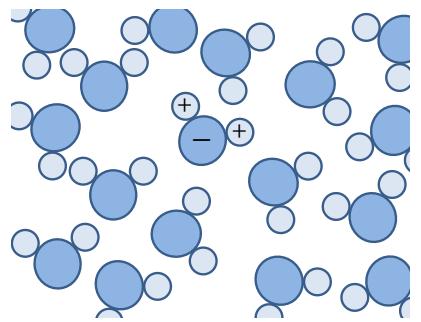
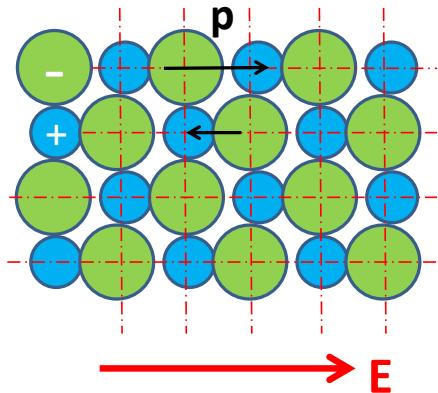
- The capacitance matrix coefficients are multiplied by ϵ_r .

- The volume density of free energy is $f(\mathbf{D}) = \frac{\mathbf{D}^2}{2\epsilon}$.

- The electrostatic pressure on a surface carrying a charge of density σ is $p = \frac{\sigma^2}{2\epsilon}$.

Chapter 5

Dielectrics: Microscopic study and effect of time-dependent fields



Introduction

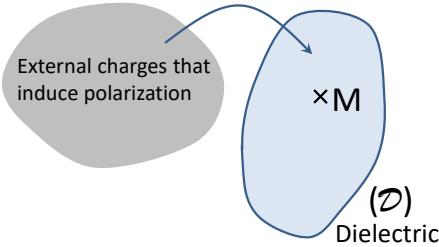
This chapter is devoted to dielectrics. Section 5.1 deals with the physical mechanism at the origin of polarization. Section 5.3 considers the response of dielectrics to time-dependent electric fields.

5.1 Microscopic study of dielectrics

In this section, we study dielectric media from the microscopic point of view. We consider media without charge ($\rho = 0$) and where the polarization \mathbf{P} is induced by an external electric field or external charges, and results in a uniform and constant polarization. The dielectric media under consideration are isotropic liquids or solids. Solid media are assumed to be crystals with a ionic or molecular lattice.

We first distinguish the effect of various fields (external field, local field, etc) on dielectrics before presenting different polarization mechanisms

5.1.1 The depolarizing field



At any point M of the dielectric medium (\mathcal{D}), there is an external field $\mathbf{E}_{\text{ext}}(M)$ due to external charges that will induce a polarization in (\mathcal{D}) . Thus, a *microscopic field* $\mathbf{e}_q(M)$ appears at any point M in response to the external field. This microscopic field takes into account the real distribution of charges within the molecules and atoms of the dielectric medium, as well as the orientation of these particles. It is therefore extremely complex and almost impossible to determine. The total microscopic electric field is simply the sum of both fields:

$$\mathbf{E}_{\text{ext}} + \mathbf{e}_q.$$

An external observer is unable to perceive all the details of this electric field and will experience only an average value, that is, the macroscopic electric field defined as a spatial average of the total microscopic field.

We therefore consider an averaging (mesoscopic) volume to define the *macroscopic field*:

$$\mathbf{E}(M) = \langle \mathbf{E}_{\text{ext}} + \mathbf{e}_q \rangle$$

It is this average field that appears in Maxwell's equations.

The polarization of the medium, that is, the appearance of dipole moments can have different origins acting simultaneously:

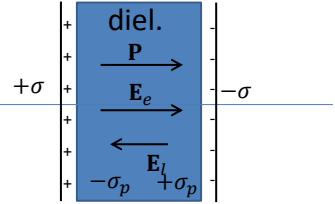
1. Atoms may undergo a distortion of the electric cloud surrounding their nuclei, that is a shift of the center of mass of the negative charge distribution with respect to the center of mass of the positive charge distribution.
2. A crystal lattice may undergo deformations, leading to the appearance of dipole moments.
3. Molecular media such as water may be polar, that is, each molecule carries a permanent dipole moment and orient itself under the action of an electric field. The medium then becomes polarized due to the orientation of existing dipoles.

The *Depolarizing electric field* denotes the field generated by the polarization \mathbf{P} . It corresponds to the second term in the expression of the spatially averaged electric field $\mathbf{E}_d \equiv \langle \mathbf{e}_q \rangle$. Hence, we can write the macroscopic field as

$$\mathbf{E}(M) = \mathbf{E}_{\text{ext}}(M) + \underbrace{\mathbf{E}_d(M)}_{\text{depolarizing field}}$$

Example:

Calculate the depolarizing field in the situation depicted on the figure: A capacitor carries a free charge of surface density $\pm\sigma$ on its plates. This free charge distribution produces a field that polarizes the LHI dielectric medium inserted between the metal plates.



The polarization \mathbf{P} of the dielectric is uniform, hence there is no volume charge polarization. In contrast, the surfaces of the dielectric in contact with the metallic plates carry a surface polarization charge $\sigma_P = \mathbf{P} \cdot \mathbf{n}$. The electric field in the dielectric due to this polarization surface charge is therefore the same as the field generated by a capacitor carrying surface charge density $\pm\sigma_P$, that is, a field of amplitude σ_P/ϵ_0 opposite to the field produced by the free charges. We conclude that the polarization of the dielectric medium generates the depolarizing field:

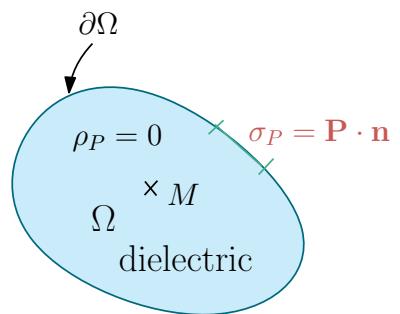
$$\mathbf{E}_d = -\frac{\mathbf{P}}{\epsilon_0}.$$

5.1.2 The local field

Consider a dielectric medium that does not contain any free charge. Let us assume that at any point M , there is a charge, for instance an electron, that belongs to an atom or a molecule of the dielectric.

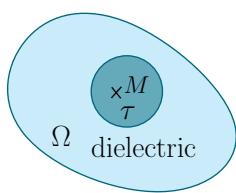
We will study the electrostatic action of the rest of the Universe, including all other (bound) charges of the dielectric and external free charges and fields, on the particle at M . The electric field generated at M by all charges except the charge at M is called the local field.

Due to our assumption that the polarization is uniform within the domain Ω , there is no volume polarization charge ($\rho_P = 0$). The depolarizing field \mathbf{E}_d induced by \mathbf{P} then corresponds to the field generated by the surface charges σ_p on $\partial\Omega$.



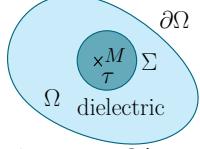
In order to define the local field acting on the particle at M , let us consider a small volume τ around M , that has typical dimension 10 to 100 times larger than the average distance between particles, atoms or molecules of typical size 10^{-10} m.

The local electric field at M includes two contributions:



- the field $\mathbf{E}_{\text{ext}}(M)$ generated at M by external charges to the dielectric domain Ω .
- the field generated by the (bound) charges within the domain Ω , from which the particle at M is omitted.

To evaluate the second contribution, we apply the superposition theorem to the domain Ω that we decompose into a small volume τ around M and the domain Ω' that includes the outer surface $\partial\Omega$ of Ω and the inner surface Σ that coincide with the outer surface of τ .



(\mathcal{D}') bounded by outer surface (\mathcal{S}) and inner surface (Σ) is a polarized domain far from M

The domain Ω' is a part of Ω and thus, it is a uniformly polarized domain but it is far from the point M . It carries no volume charge polarization ($\rho_P = 0$) but it possibly carries a surface charge polarization of density $\sigma_P = \mathbf{P} \cdot \mathbf{n}$ on both surfaces $\partial\Omega$ and Σ . The polarization charge σ_P on Σ (resp. on $\partial\Omega$) is at the origin of an electric field at M : $\mathbf{E}_\Sigma(M)$ (resp. $\mathbf{E}_{\partial\Omega}(M)$). The local field at M is then obtained by gathering all contributions:

$$\mathbf{E}_l(M) = \mathbf{E}_{\text{ext}}(M) + \mathbf{E}_{\partial\Omega}(M) + \mathbf{E}_\Sigma(M) + \mathbf{E}_\tau(M).$$

• EVALUATION OF $\mathbf{E}_{\partial\Omega}(M)$

The field at M generated by the polarization charges on (S) can be evaluated from the general formula

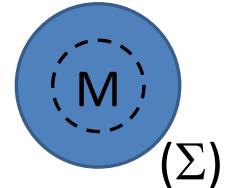
$$\mathbf{E}_{\partial\Omega}(M) = \oint_{\partial\Omega} \frac{\sigma_P(\mathbf{r}')(\mathbf{r}' - \mathbf{r})}{4\pi\epsilon_0|\mathbf{r}' - \mathbf{r}|^3} dS(\mathbf{r}') = \oint_{\partial\Omega} \frac{\mathbf{P} \cdot \mathbf{n}(\mathbf{r}')(\mathbf{r}' - \mathbf{r})}{4\pi\epsilon_0|\mathbf{r}' - \mathbf{r}|^3} dS(\mathbf{r}')$$

However, we will not need to evaluate this integral explicitly. Indeed, since the polarization is uniform in domain Ω , $\nabla \cdot \mathbf{P} = 0$ and there is no volume polarization charge that contribute to the depolarizing field, thus, the depolarizing field coincides with the field evaluated from the surface polarization charge on $\partial\Omega$:

$$\mathbf{E}_{\partial\Omega}(M) = \mathbf{E}_d(M).$$

• EVALUATION OF $\mathbf{E}_\tau(M)$

We evaluate now the electric field $\mathbf{E}_\tau(M)$ generated by the charged particles inside the volume τ (bounded by Σ), from which the particle at M is omitted. Note that the surface polarization charge on Σ contribute to $\mathbf{E}_\Sigma(M)$, not to $\mathbf{E}_\tau(M)$.



We will limit our calculation to the two following situations where

- the medium is a polar liquid. In this case the medium is isotropic. We chose the volume τ to be a sphere centered at M . At the center of the sphere, the field generated by all the charges within τ is zero because of spherical symmetry.
- the medium is an isotropic cubic crystal. In this case we chose the surface Σ to be a cube centered at M , whose faces are parallel to the faces of the cubic lattice of the crystal. We admit that due to the cubic symmetries of the crystal lattice, the field at M is again zero.

Thus, in both cases the charges within the volume τ do not contribute to the local field

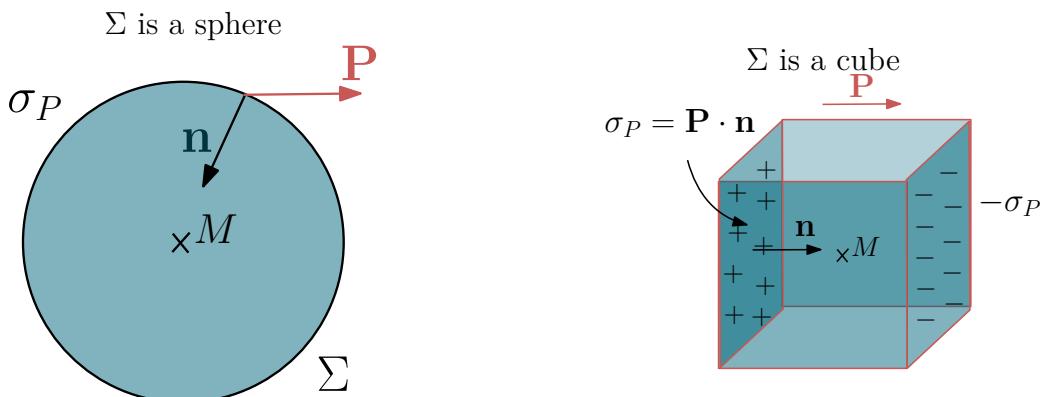
$$\mathbf{E}_\tau(M) = \mathbf{0}$$

- EVALUATION OF $\mathbf{E}_\Sigma(M)$

In the two cases under consideration, that is, when Σ is a spherical surface or a cube centered at M , we will evaluate the field $\mathbf{E}_\Sigma(M)$ generated at M by the polarization charge lying on the surface Σ .

Since the polarization is uniform, the surface polarization charge on the spherical surface is $\sigma_P = \mathbf{P} \cdot \mathbf{n}$, where \mathbf{n} is the normal to the sphere pointing outward, that is, toward M . Hence $\sigma_P = P \cos \theta$, where θ denotes the angle of spherical coordinates with the z -axis parallel to \mathbf{P} . In that case, we have shown in an exercise (see chapter 2) that the field at M generated by this surface charge distribution reads

$$\mathbf{E}_\Sigma(M) = \frac{\mathbf{P}}{3\epsilon_0}$$



If the surface Σ is a cube chosen so as to have four faces parallel to the uniform polarization, these four faces have a normal that is perpendicular to \mathbf{P} and hence do not carry any surface polarization charge. The two other faces carry a surface polarization charge $\sigma_P = \pm P$. We have already calculated (supplementary exercise in lecture 2) the electric field at the center of the cube and found

$$\mathbf{E}_\Sigma(M) = \frac{\mathbf{P}}{3\epsilon_0}.$$

- FINAL RESULT FOR THE LOCAL FIELD

Gathering all contributions to the local field, we can write

$$\mathbf{E}_l(M) = \underbrace{\mathbf{E}_{\text{ext}}(M)}_{\text{local}} + \underbrace{\mathbf{E}_d(M)}_{\text{external}} + \underbrace{\frac{\mathbf{P}}{3\epsilon_0}}_{\text{depolarizing}} + \mathbf{0}$$

Since the sum of the first two terms on the right hand side correspond to the definition of the macroscopic field, we finally obtain:

$$\mathbf{E}_l(M) = \underbrace{\mathbf{E}(M)}_{\text{macroscopic field}} + \underbrace{\frac{\mathbf{P}}{3\epsilon_0}}_{\text{Corrective field}}$$

We note that the local field that applies on a particle is different from the macroscopic field. This means that the motion of particle at M should be described from the action of the local field acting on the particle, that is, the force entering in Newton's law reads $\mathbf{f} = q\mathbf{E}_l(M)$.

The corrective field that enters in the expression for the local field is important for polar liquids and crystals. It is less relevant for dilute media. Let us evaluate orders of magnitude to justify this assertion: For a LHI medium, the polarization is linked to the electric field: $\mathbf{P} = \epsilon_0\chi\mathbf{E}$. This allows us to express the local field as a function of the macroscopic field:

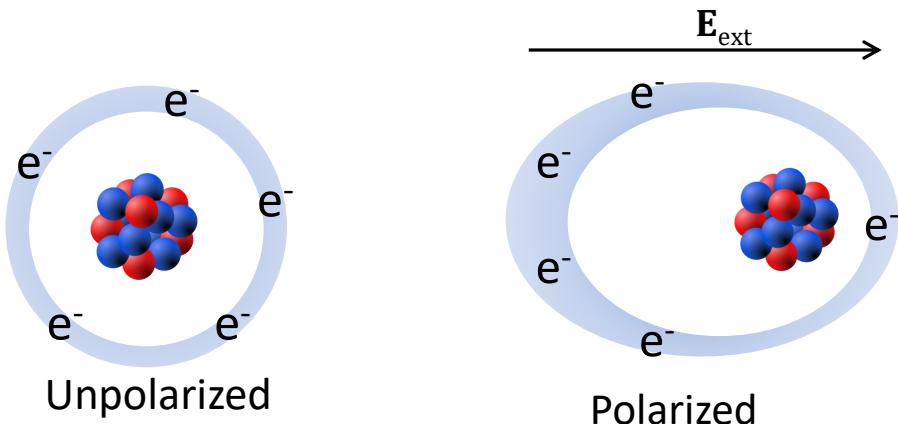
$$\mathbf{P} = \epsilon_0\chi\mathbf{E} \Rightarrow \mathbf{E}_l = \left(1 + \frac{\chi}{3}\right)\mathbf{E}, \quad \chi = \epsilon_r - 1.$$

Clearly, the local field differs from \mathbf{E} if the susceptibility χ cannot be neglected, that is if the relative permittivity significantly differs from 1 ($\epsilon_r \neq 1$), which is precisely the case for dense media: liquids and crystals. For dilute media (gases), $\chi \ll 1 \Rightarrow \mathbf{E}_l(M) \sim \mathbf{E}(M)$.

5.2 Polarization mechanisms

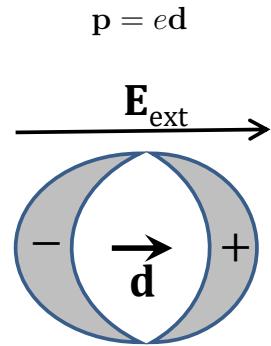
5.2.1 Electronic polarization

Under the action of the local field \mathbf{E}_l , the electron cloud in an atom undergoes a distortion that is accompanied by the generation of a dipole moment.



Consider a simple model of atom that has spherical symmetry for the constituting charge distributions. Within a sphere of radius R , the protons of charge $+e$ have a volume charge density $\rho^+ = 3e/4\pi R^3$; the electrons of charge $-e$ have a charge density $\rho^- = -3e/4\pi R^3$.

Under the action of the local field, the electron cloud moves to the left by a small quantity $d \ll R$, generating a dipole of moment $\mathbf{p} = e\mathbf{d}$. The positive charge undergoes the actions of the local field \mathbf{E}_l and the action of the electric field \mathbf{E}_- resulting from the displacement of the negative charge distribution. This displacement constitute a cavity problem. We know that applying the Gauss theorem to a uniformly charged sphere and the superposition theorem allows us to find



$$\mathbf{E}_- = \frac{\rho^- \mathbf{d}}{3\epsilon_0}$$

The equilibrium of the positive charge thus reads

$$\mathbf{E}_l + \mathbf{E}_- = 0$$

Hence,

$$\mathbf{E}_l = -\frac{\rho^-}{3\epsilon_0} \mathbf{d} = +\frac{ed}{4\pi\epsilon_0 R^3} \mathbf{u}_x = \frac{\mathbf{p}}{4\pi\epsilon_0 R^3}.$$

We define the *polarizability* α of the atom by the linear relation between the dipole moment and the local field

$$\mathbf{p} = \epsilon_0 \alpha \mathbf{E}_l.$$

Thus, the expression for the dipole moment $\mathbf{p} = 4\pi\epsilon_0 R^3 \mathbf{E}_l$ allows us to identify the polarizability

$$\alpha = 4\pi R^3$$

[α] : volume

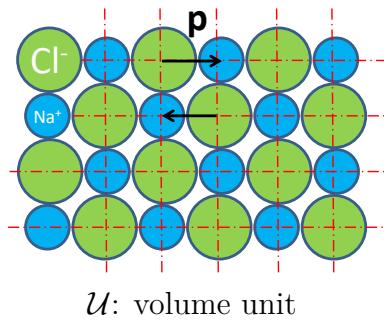
Note that the polarizability has the dimensions of a volume. It roughly corresponds to the volume of the atom.

5.2.2 Ionic polarization

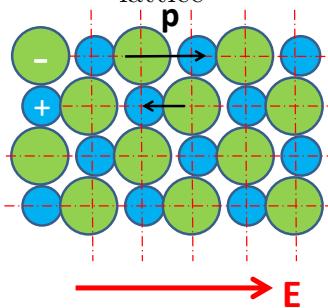
In the absence of an external field, the ions in a crystal that does not carry a permanent dipole moment have charges q_i and positions \mathbf{r}_i . The polarization in the crystal is expressed for the volume unit \mathcal{U} and is zero by definition:

$$\mathbf{P} = \sum_{i \in \mathcal{U}} e(\mathbf{r}_i^+ - \mathbf{r}_i^-) = \mathbf{0}$$

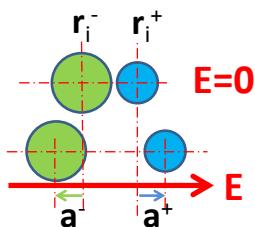
Non-permanently polarized crystal



Field $\mathbf{E}_l \rightarrow$ distortion of the lattice



Under the action of a local field \mathbf{E}_l , some ions undergo a displacement by a small quantity \mathbf{a}_i with respect to their previous position. A polarization appears, expressed as



$$\begin{aligned} \mathbf{P} &= \sum_{i \in U} e(\mathbf{r}_i^+ + \mathbf{a}^+ - \mathbf{r}_i^- - \mathbf{a}^-) \\ &= \sum_{i \in U} e(\mathbf{a}_i^+ - \mathbf{a}_i^-) = ne \underbrace{(\mathbf{a}^+ - \mathbf{a}^-)}_{\text{proportional to } \mathbf{E}_l}, \end{aligned}$$

where we have introduced the condition that an absence of electric field corresponds to a non-polarized crystal. We implicitly assumed that the crystal has two types of atoms, for instance sodium chloride, so that all positive (resp. negative) charges undergo a displacement by the same quantity \mathbf{a}^+ (resp. \mathbf{a}^-) under the action of the field. Finally, we assumed that the crystal contains n atoms per unit volume.

In order to express the polarizability, we consider the equilibrium of a cation, given by Newton's law, where the electric force due to the local field is balanced by the restoring force proportional to the relative displacement of cations with respect to anions¹. Along the x -axis, we find:

$$eE_l - Y|\mathbf{r}_i^+ - \mathbf{r}_i^-||\mathbf{a}^+ - \mathbf{a}^-| = 0,$$

where Y denotes the Young modulus of the crystal. This allows us to rewrite the polarization

$$\mathbf{P} = ne(\mathbf{a}^+ - \mathbf{a}^-) = n \frac{e^2}{Y|\mathbf{r}_i^+ - \mathbf{r}_i^-|} \mathbf{E}_l.$$

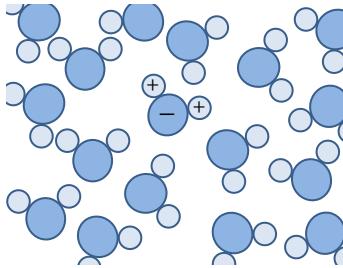
Finally, we can identify the polarizability:

$$\mathbf{P} = n\epsilon_0\alpha\mathbf{E}_l \Rightarrow \boxed{\alpha = \frac{e^2}{\epsilon_0 Y |\mathbf{r}_i^+ - \mathbf{r}_i^-|}}.$$

¹We can write this force as the product of the stress by the surface area $|\mathbf{r}_i^+ - \mathbf{r}_i^-|^2$ and using Hooke's law for the stress $\sigma = Y\epsilon$, where Y denotes the Young modulus of the crystal and $\epsilon = \frac{|\mathbf{a}^+ - \mathbf{a}^-|}{|\mathbf{r}_i^+ - \mathbf{r}_i^-|}$ denotes the strain or proportional deformation.

5.2.3 Polarization by dipolar orientation

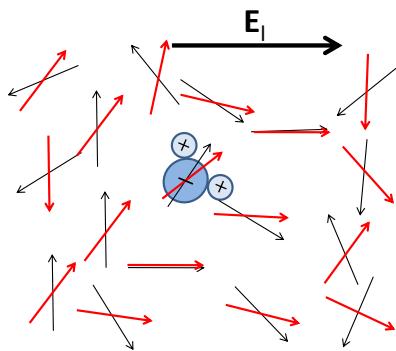
In this section, we consider dielectric media that contains polar molecules, such as water. The molecules carry a permanent dipole moment \mathbf{p}_{0i} of magnitude p_0 . Under the action of an external electric field, the dipole moments, and hence the molecules, align with the field and the medium acquires a polarization. It is this polarization by molecular orientation that is investigated in this section.



medium that contains polar molecules (H_2O)
 $|\mathbf{p}_0| = p_0$

In the absence of any external electric field, dipoles (molecules) are randomly oriented. The polarization of the medium, that is, the sum of the dipole moments over a unit volume is zero even if each molecule is permanently polarized

$$\mathbf{P} = \sum_{i \in U} \mathbf{p}_{0i} = \mathbf{0}.$$



In the presence of a local field \mathbf{E}_l , dipoles tend to align parallel to the field $\mathbf{p}_0 \parallel \mathbf{E}_l$. However, thermal agitation tends to disorient the dipoles, hence, only partial (average) alignment with the local field is obtained for each dipole. In result, summing over all the dipoles in the volume unit leads to a polarization \mathbf{P}

$$0 \leq \mathbf{P} \leq n\mathbf{p}_0.$$

We will evaluate the polarization by modeling the competition between alignment of dipoles parallel to the field and thermal agitation.

Thermal agitation is governed by Boltzmann's law, which states that the probability $d\mathcal{P}$ to observe a particle at temperature T with an energy in the range $[W, W + dW]$ reads

$$d\mathcal{P} = \mathcal{P}_0 e^{-\frac{W}{k_B T}} dW$$

where $k_B = 1.38 \times 10^{-23} \text{ J K}^{-1}$ denotes the Boltzmann constant and W is the energy of the particle. In the present case, W is the energy of the molecular dipole in the local field, hence

$$W = -\mathbf{p}_0 \cdot \mathbf{E}_l = -p_0 E_l \cos \theta, \quad 0 < \theta < \pi.$$

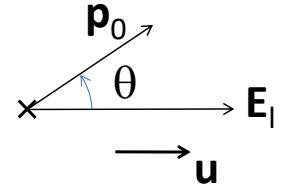
The constant \mathcal{P}_0 denotes a normalization value, such that the sum of probabilities be equal to one when W covers all possible energies.

For the convenience of notations, let us define $W_0 = p_0 E_l$, which represents the energy bound for a dipole in the local field: $-W_0 \leq W \leq +W_0$.

The contribution of a single molecule to the polarization is $p_0 \cos \theta \mathbf{u} d\mathcal{P}$, where \mathbf{u} denotes a unit vector along the local field, and θ denotes the orientation of a molecule of energy W , that is, the angle between its dipole moment and the local field \mathbf{E}_l .

The calculation of the polarization \mathbf{P} is then performed by integration over all energies, or equivalently over all possible molecular orientations

$$\mathbf{P} = \mathbf{u} n p_0 \frac{\int_{-W_0}^{W_0} \cos \theta \mathcal{P}_0 e^{-\frac{W}{kT}} dW}{\int_{-W_0}^{W_0} \mathcal{P}_0 e^{-\frac{W}{kT}} dW}.$$



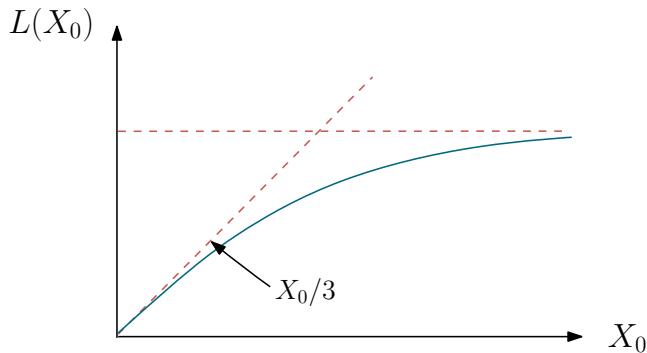
The calculation can be performed analytically by using the change variable

$$\cos \theta = -\frac{W}{W_0}, \quad X = -\frac{W}{kT},$$

and we find

$$\mathbf{P} = np_0 \mathbf{u} \frac{kT}{W_0} \frac{\int_{-X_0}^{X_0} X e^X dX}{\int_{-X_0}^{X_0} e^X dX} = np_0 \mathbf{u} \underbrace{\left(\frac{1}{\tanh X_0} - \frac{1}{X_0} \right)}_{\mathcal{L}(X_0)}$$

where $X_0 = +\frac{W_0}{kT}$. The function $\mathcal{L}(X) = 1/\tanh(X) - 1/X$ denotes the Langevin function and is plotted on the figure below.



- When X_0 is large, that is for low temperatures T , the Langevin function tends to saturation ($\mathcal{L}(X_0) \rightarrow 1$). In this case, all molecules tend to be aligned with the field and

$$\mathbf{P} = np_0 \mathbf{u}.$$

- When X_0 is small, that is for high temperatures T or when the field E_l is weak, the Langevin function $\mathcal{L}(X_0) \simeq X_0/3$. In this case, the expression for the polarization as a function of the local field reads

$$\mathbf{P} = np_0 \mathcal{L}(X_0) \frac{\mathbf{E}_l}{|\mathbf{E}_l|} \Rightarrow \mathbf{P} = np_0 \frac{p_0}{3kT} \mathbf{E}_l,$$

which allows us to identify the polarizability for molecular orientation

$$\alpha = \frac{p_0^2}{3\epsilon_0 kT} \quad (\text{weak fields})$$

The medium acquires an average macroscopic polarization that depends on its temperature. For molecular media such as water, all polarization sources contribute. Thus, the electronic polarizability must be also taken into account in the total polarization

$$\mathbf{P} = \epsilon_0 n \left(\alpha_e + \frac{p_0^2}{3\epsilon_0 kT} \right) \mathbf{E}_l$$

5.2.4 Clausius-Mossotti's relation

Several physical mechanisms at the origin of polarization were identified in previous sections. They allowed us to find expressions for the polarizability α , that links the macroscopic polarization density \mathbf{P} to the local field \mathbf{E}_l :

$$\mathbf{P} = \epsilon_0 n \alpha \mathbf{E}_l.$$

In the previous chapter, we have defined the dielectric constant or relative permittivity ϵ_r , linking the macroscopic polarization and the electric field \mathbf{E} :

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} = \epsilon_0 (\epsilon_r - 1) \mathbf{E}.$$

Clausius-Mossotti's relation, also called Lorentz-Lorenz's relation, links the polarizability of a dielectric medium and its dielectric constant, or susceptibility $\chi = \epsilon_r - 1$.

α	χ	ϵ_r
polarizability	susceptibility	dielectric constant
$\mathbf{P} = n\epsilon_0 \alpha \mathbf{E}_l$	$\mathbf{P} = \epsilon_0 \chi \mathbf{E}$	$\epsilon_r = 1 + \chi$

In a linear, homogeneous and isotropic medium, the local field is expressed as a function of the external field and the polarization density:

$$\mathbf{E}_l = \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0}.$$

Eliminating the local field in the expression for the polarization density, we find

$$\mathbf{P} = n\epsilon_0\alpha\mathbf{E}_l = n\epsilon_0\alpha \left(\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right) \Rightarrow \left(1 - \frac{n\alpha}{3} \right) \mathbf{P} = n\epsilon_0\alpha\mathbf{E}$$

Thus, introducing the relation $\mathbf{P} = \epsilon_0(\epsilon_r - 1)\mathbf{E}$, we obtain $n\epsilon_0\alpha = \epsilon_0(\epsilon_r - 1) \left(1 - \frac{n\alpha}{3} \right)$ which can be rewritten in the form of Clausius-Mossotti's relation:

$$\frac{n\alpha}{3} = \frac{\epsilon_r - 1}{\epsilon_r + 2}$$

Clausius-Mossotti relation

Equivalently, we can write the expression for the susceptibility $\chi = \epsilon_r - 1$:

$$\chi = \frac{n\alpha}{1 - \frac{n\alpha}{3}}.$$

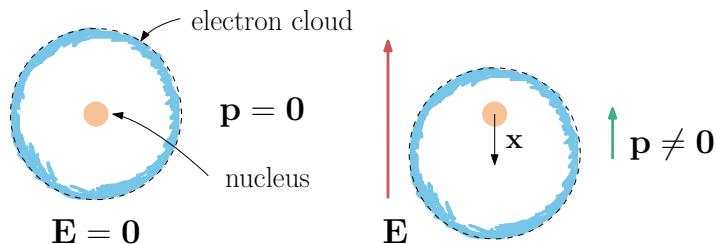
The dielectric constant is therefore a function of the temperature, via the polarizability α , and of the pressure, via the density of molecules per volume unit n .

5.3 Dielectrics in time-dependent fields

5.3.1 The Lorentz model for the classical atom (1878)

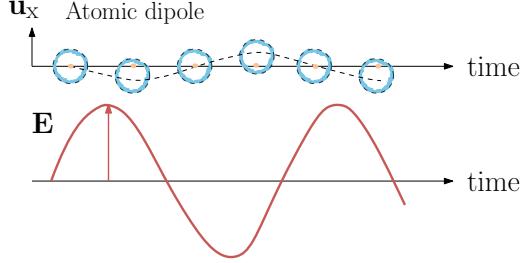
We will establish a model for the frequency-dependence of the dielectric constant.

Let us recall that in an atom, under the presence of an electric field, the electron cloud orbiting around the nucleus will be shifted from its equilibrium position by a quantity \mathbf{x} . In consequence an atomic electric dipole \mathbf{p} develops. Here, we neglect the movement of the nucleus in the electric field due to its much larger mass compared to that of an electron.



In the presence of an electromagnetic wave, the electric field oscillates periodically, and therefore, so does the induced dipole. We expect that if the frequency ω of the wave is very low, the dipole will instantaneously adapt to the electric field and oscillate in phase with the latter, such that $\lim_{\omega \rightarrow 0} \epsilon(\omega) = \epsilon$, where ϵ is the static permittivity. In contrast, if the frequency is too high, the inertia of the electron cloud will make it impossible for the electrons to follow

the rapid oscillations of the electric field, and the medium will become transparent. We then expect the permittivity to tend asymptotically to ϵ_0 as $\omega \rightarrow \infty$, that is, $\lim_{\omega \rightarrow \infty} \epsilon(\omega) = \epsilon_0$ and waves of very high frequency will propagate as if they were in vacuum.



The Lorentz model takes into account two crucial ingredients: firstly, since an oscillating dipole generates in fact an electric current, it creates its own electromagnetic wave. This means that a dipole driven at a frequency ω will lose part of its mechanical energy by emitting radiation at a frequency close to ω . Secondly, we know from experiments that electrons in atoms have discrete energy levels and as such, they absorb or emit light at precise resonant frequencies, a fact that arises from the quantum behavior of matter at atomic scales. The Lorentz model is a classical model that incorporates these two effects: the movement of an electron around the nucleus is modeled by a damped harmonic oscillator with a natural frequency ω_0 . The equation of motion reads

$$m_e \frac{d^2 \mathbf{x}}{dt^2} = -e \mathbf{E} - m_e \gamma \frac{d\mathbf{x}}{dt} - m_e \omega_0^2 \mathbf{x}$$

where \mathbf{x} is the displacement of the electron with respect to the equilibrium position around the nucleus and γ is a damping factor that accounts for the fact that energy is absorbed from the incident electromagnetic wave and re-emitted by the electrons (in an atom γ is typically of the order of 10^8 s^{-1}). Here we neglected the magnetic component of the Lorentz force since, for a plane electromagnetic wave and for velocities much smaller than the speed of light,

$$| -e \mathbf{v} \times \mathbf{B} | \leq e |\mathbf{v}| |\mathbf{B}| = e \frac{|\mathbf{v}|}{c} |\mathbf{E}| \ll e |\mathbf{E}|$$

Assuming now a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$ with a wavelength much longer than the oscillation amplitude of the atomic dipole, we may consider the field to be spatially uniform at the position of the dipole and look for a stationary solution in the form of a forced oscillator $\underline{\mathbf{x}} = \underline{\mathbf{x}}_0 e^{-i\omega t}$, so we have

$$-m_e \omega^2 \underline{\mathbf{x}} = -e \underline{\mathbf{E}} + i\omega m_e \gamma \underline{\mathbf{x}} - m_e \omega_0^2 \underline{\mathbf{x}}.$$

Then,

$$\underline{\mathbf{x}} = \frac{e}{m_e \omega^2 - \omega_0^2 + i\gamma\omega} \frac{\underline{\mathbf{E}}}{\underline{\mathbf{E}}_0}. \quad (5.1)$$

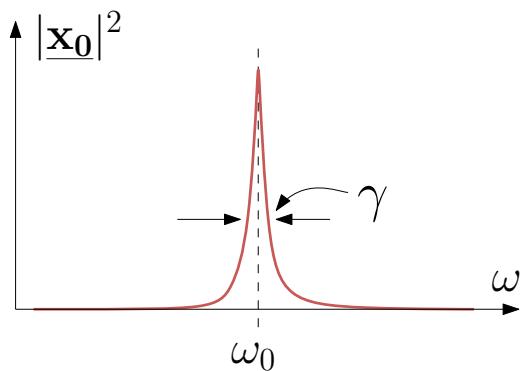
The squared amplitude of the electron oscillation then reads

$$|\underline{\mathbf{x}}_0|^2 = \frac{e^2}{m_e^2} \frac{|\underline{\mathbf{E}}_0|^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2 \omega^2}.$$

and exhibits a resonance at $\omega_{\text{res}} = \sqrt{\omega_0^2 - \gamma^2/2} \simeq \omega_0$ for which the amplitude

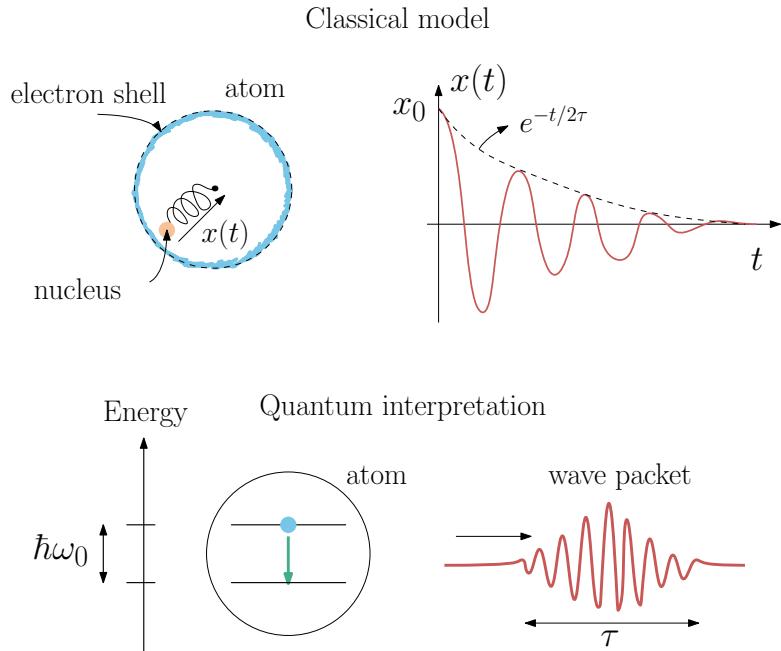
$$|\underline{x}_0(\omega_{\text{res}})| = \frac{e}{m_e} \frac{|\underline{E}_0|}{\gamma \sqrt{\omega_0^2 - \gamma^2/4}} \simeq \frac{e}{m_e} \frac{|\underline{E}_0|}{\gamma \omega_0}$$

is very large (this would correspond to a transition in quantum mechanics). The width of the resonance is given by γ which, for an isolated atom ($\gamma = 10^8 \text{ s}^{-1}$) and a visible electromagnetic wave ($\omega_0 \sim 3 \times 10^{15} \text{ s}^{-1}$) is such that $\gamma/\omega_0 \sim 10^{-8}$. In that case the amplitude as a function of frequency has approximately a Cauchy-Lorentzian shape as shown in the figure below



Remarks

- The damping frequency γ is related to the typical time $\tau = 1/\gamma$ that it takes for the amplitude of an oscillator to decay by a factor of e , and it is related to the energy lost in emitted radiation by the oscillating dipole. In quantum mechanics, τ has a different interpretation: an electron in an excited state will spontaneously decay to the lowest energy state within a time τ , emitting a wave packet of frequency ω_0 and of typical duration τ . The frequency ω_0 of the emitted wave is related to the energy difference ΔE between the excited and lowest energy state, by $\Delta E = \hbar\omega_0$, where \hbar is the reduced Planck constant.



- This simple model, although representing an atom, can be extended to describe more complex systems such as molecules (whose vibrational modes are quantized as well and therefore exhibit resonant frequencies) or crystals. As such, gases, liquids or solid media can be modeled by an assembly of harmonic oscillators. Typically, the linewidth γ of the resonances in matter is dominated by collisions between the atoms and not by spontaneous emission, as in the case of isolated atoms.

• THE COMPLEX PERMITTIVITY

If n is the density of atomic (or molecular) dipoles in the medium, each one having Z electrons and a single resonance ω_0 , the polarization vector reads

$$\underline{\mathbf{P}} = n(-Ze\underline{\mathbf{x}}) = \frac{-nZe^2}{m_e} \frac{\underline{\mathbf{E}}}{\omega^2 - \omega_0^2 + i\gamma\omega} = (\underline{\epsilon}(\omega) - \epsilon_0)\underline{\mathbf{E}},$$

so that the complex permittivity is given by

$$\underline{\epsilon}(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \right), \quad (5.2)$$

where $\omega_p = \sqrt{Zne^2/m_e\epsilon_0}$ is the plasma frequency. Separating explicitly its real and imaginary parts gives

$$\epsilon(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} + i \frac{\omega_p^2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \right). \quad (5.3)$$

Note that in the limit of very high frequencies, the medium becomes transparent, $\lim_{\omega \rightarrow \infty} \epsilon(\omega) = \epsilon_0$, whereas for very low frequencies we get an expression for the static permittivity,

$$\lim_{\omega \rightarrow 0} \epsilon(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2}{\omega_0^2} \right) = \epsilon.$$

• THE COMPLEX REFRACTIVE INDEX

Finally, the complex index of refraction n , which fully characterizes the propagation of a wave in a medium, is complex whenever $\epsilon(\omega)$ has a non-negligible imaginary part (i.e. close to resonant frequencies). In this case, it writes

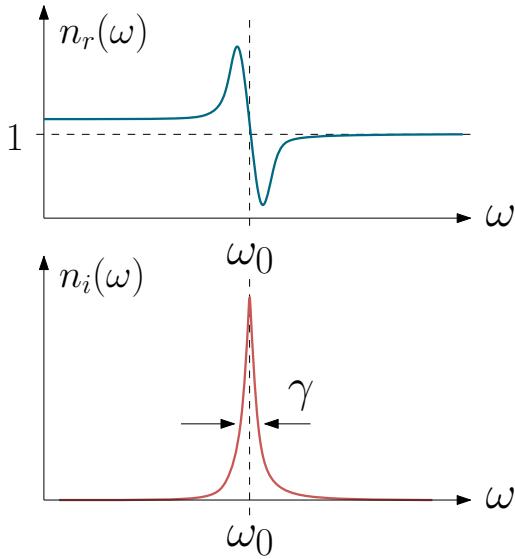
$$n(\omega) = \sqrt{\frac{\epsilon(\omega)}{\epsilon_0}} = n_R(\omega) + i n_I(\omega).$$

For a diluted medium, one has $\omega_p \ll \omega_0$. We can then use the expansion

$$\sqrt{\frac{\epsilon(\omega)}{\epsilon_0}} = \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \right)^{\frac{1}{2}} \approx 1 + \frac{\omega_p^2}{2(\omega_0^2 - \omega^2 - i\gamma\omega)}$$

and so,

$$n(\omega) \approx \underbrace{1 + \frac{1}{2} \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}_{n_R} + i \underbrace{\frac{1}{2} \frac{\omega_p^2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}_{n_I}. \quad (5.4)$$



Remarks

- Away from any resonance, $n_I \ll n_R$ so that absorption may be neglected and the refractive index may be considered as a pure real number, $n \approx n_R$.

- The very large change of index of refraction n_R close to a resonance will cause a large dispersion, i.e., different wavelengths travel at very different speeds near the resonance.
- Note that for frequencies slightly above a very strong resonance, the real part of the refractive index could be less than one. This does not mean, however, than an electromagnetic wave centered around this frequency will travel faster than c . As was shown in the optics course, any physical wave is represented by a wave packet, that is, a superposition of plane waves. The speed at which such a wave packet moves is not given by $\omega/k(\omega)$ but instead by $\partial\omega/\partial k$, also called group velocity.
- Finally, in real materials there are several types of resonances distributed in the whole electromagnetic spectrum (vibrations fall in the infrared, electronic resonances in atoms typically in the UV or visible range, etc). A realistic model including more than one resonant frequency reads

$$\underline{n}^2(\omega) = 1 + \omega_p^2 \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - i\gamma_j\omega},$$

where f_j , called the oscillator strength, represents the relative weight of the frequency ω_j which can be properly determined within the framework of quantum mechanics.

5.3.2 Orientation polarization in the sinusoidal regime

In this section, we consider a material with molecules carrying a permanent dipole moment \mathbf{p}_0 (e.g. water) that can freely rotate to align with an electric field. It is the response of these dipoles to a sinusoidal time-dependent electric field that we will investigate.

Consider first a constant applied electric field \mathbf{E}_0 . We have shown that the molecules respond to \mathbf{E}_0 via the local field \mathbf{E}_l . In the presence of a local field \mathbf{E}_l , each molecule becomes polarized with a dipole moment that is expressed by its average

$$\mathbf{p} = \epsilon_0 \alpha \mathbf{E}_l,$$

with

$$\alpha = \frac{p_0^2}{3\epsilon_0 k T}$$

We have established that in the permanent regime, the dielectric susceptibility of the medium reads

$$\chi_0 = \frac{n\alpha}{1 - \frac{n\alpha}{3}}.$$

Now assume that the external applied field \mathbf{E}_0 suddenly goes back to zero. The material then returns to the non polarized stationary regime but this does not occur instantaneously: the transient depolarization of the medium, from a polarization \mathbf{P}_0 , follows an exponential law:

$$\mathbf{P} = \mathbf{P}_0 e^{-t/\tau},$$

which corresponds to the solution to the differential equation

$$\frac{d\mathbf{P}}{dt} + \frac{\mathbf{P}}{\tau} = \mathbf{0}. \quad (5.5)$$

The system becomes non polarized $\mathbf{E} = 0, \mathbf{P} = 0$ after a delay of a few τ , which represents the relaxation time. If the electric field is suddenly established to a non-zero constant value \mathbf{E}_0 , the polarization will similarly not reach immediately its permanent regime value $\mathbf{P}_0 = \epsilon_0 \chi_0 \mathbf{E}_0$, but will grow with a characteristic time equal to the relaxation time τ , the order of magnitude of which is a few tens of picoseconds for water molecules (30×10^{-12} s).

We now consider the time dependent regime. The electric field $\mathbf{E}(t)$ is a time dependent function and the polarization $\mathbf{P}(t)$ obeys the ordinary differential equation

$$\tau \frac{d\mathbf{P}}{dt} + \mathbf{P} = \epsilon_0 \chi_0 \mathbf{E}.$$

We consider the particular case of a sinusoidal electric field² of frequency ω . Hence, in complex notations, $\mathbf{E}(t) = \mathcal{E} \exp(-i\omega t)$, where \mathcal{E} is the vector of amplitudes for the components of the electric field. We look for a solution in the form $\mathbf{P}(t) = \mathcal{P} \exp(-i\omega t)$, where \mathcal{P} is a complex valued amplitude vector. Introducing it in Eq. 5.5, we find $d\mathbf{P}(t)/dt = -i\omega \mathbf{P}(t)$, leading to

$$\mathcal{P}(\omega) = \epsilon_0 \frac{\chi_0}{1 - i\omega\tau} \mathcal{E}$$

Now introduce the complex value susceptibility $\chi(\omega)$ from the relation

$$\mathcal{P}(\omega) = \epsilon_0 \chi(\omega) \mathcal{E}(\omega).$$

Note that the above relation is the similar to that in the permanent regime but now the linear dependence of the polarization vector upon the electric field vector is not written for the time dependent components of the field and polarization but for each frequency dependent (complex) amplitude³. From this relation, we can identify the complex susceptibility

$$\chi(\omega) = \frac{\chi_0}{1 - i\omega\tau},$$

and the complex dielectric constant, or relative permittivity, $\epsilon_r(\omega) = 1 + \chi(\omega)$:

$$\epsilon_r(\omega) = 1 + \frac{\chi_0}{1 - i\omega\tau},$$

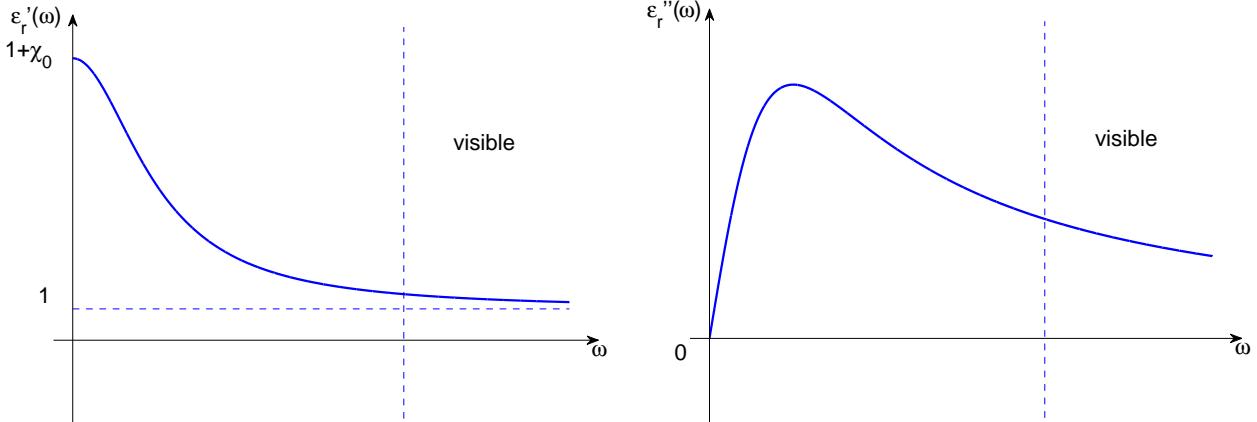
whose real and imaginary parts $\epsilon'_r(\omega)$ and $\epsilon''_r(\omega)$ are expressed as

$$\epsilon'_r(\omega) = 1 + \frac{\chi_0}{1 + \omega^2\tau^2} \quad \epsilon''_r(\omega) = \chi_0 \frac{\omega\tau}{1 + \omega^2\tau^2}.$$

²For a more general time dependence of the electric field, we would perform a Fourier decomposition into sinusoidal functions of frequency ω and use the superposition theorem to obtain the complete polarization.

³The linear relation is therefore valid for the Fourier components of the electric field and polarization.

The real part of the permittivity is related to the refractive index and tells us about the ability of electromagnetic fields (waves) to propagate in the material. The imaginary part of the permittivity is related to the absorption of electromagnetic fields by the material (see next section).



The real and imaginary parts of the permittivity vary with frequency as indicated on the figures. At high frequency ($\omega \gg 10^{12}$ Hz), the real permittivity tends to one. This means that the molecules in the material cannot follow the fast variation of the electric field because of their inertia. The orientation polarization becomes negligible. It is however still possible to obtain a non zero electronic polarization since the response time of electrons is much smaller than the response time of molecules.

At low frequency $\omega \ll 1/\tau$, the permittivity is close to the $1 + \chi_0$, which corresponds to the highest permanent polarization. For example for water molecules, $\chi_0 \simeq 80$, $\epsilon_r'(0) = 81$. The imaginary part of the permittivity exhibits a maximum $\epsilon_r''(\omega_0) = \chi_0/2$ at $\omega_0 = 1/\tau$.

5.3.3 The complex permittivity - dielectric heating

In the previous sections, we studied several models describing the response of linear homogeneous and isotropic dielectric media to a sinusoidal electric field. These media were characterized by a complex valued permittivity.

$$\mathbf{D} = \epsilon_0 \epsilon_r(\omega) \mathbf{E} \quad \text{with} \quad \epsilon_r(\omega) = \epsilon_r'(\omega) + i \epsilon_r''(\omega)$$

In this section, we will write the complex permittivity as

$$\epsilon_r(\omega) = \epsilon_r^{(0)}(\omega) e^{i\delta}$$

where δ , called the loss angle, is such that $\tan \delta = \epsilon_r''(\omega)/\epsilon_r'(\omega)$.

Using complex notations, we can write for a sinusoidal electric field:

$$\mathbf{E} = \mathcal{E}_0 e^{-i\omega t}$$

where $\mathcal{E}_0 = \mathcal{E}_0 \mathbf{u}$ is the product of the electric field amplitude \mathcal{E}_0 by the unit vector along the polarization direction of the electric field. Similarly, the displacement vector reads

$$\mathbf{D} = \mathcal{D}(\omega) e^{-i\omega t}$$

For LHI media, the displacement field and the electric field satisfy

$$\mathcal{D}(\omega) = \epsilon_0 \epsilon_r(\omega) \mathcal{E}(\omega).$$

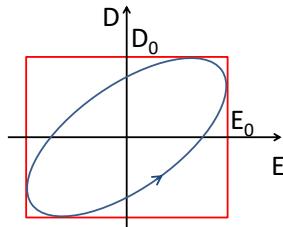
Thus, the fields \mathbf{E} and \mathbf{D} are not in phase and their complex amplitudes satisfy

$$\mathcal{D}(\omega) = \epsilon_0 \epsilon_r^{(0)}(\omega) \mathcal{E}_0 e^{i\delta}$$

Note that complex notations are used only by convenience but the same results is of course obtained using real notations:

$$E = \mathcal{E}_0 \cos(\omega t) \quad \text{and} \quad D = \underbrace{\epsilon_0 \epsilon_r^{(0)} \mathcal{E}_0}_{\mathcal{D}_0} \cos(\omega t - \delta)$$

If we plot the medium response (the displacement field) as a function of the electric field, that is, for any time t , we plot the curve $|\mathbf{D}(t)|$ as a function of $|\mathbf{E}(t)|$, we obtain an ellipse, inscribed in a rectangle of width \mathcal{E}_0 and height \mathcal{D}_0 .



$$\delta > 0$$

This ellipse gives a graphical representation of the energy losses. Indeed the work required for the displacement vector to vary from \mathbf{D} to $\mathbf{D} + d\mathbf{D}$ during dt corresponds to the free energy change, hence, per unit volume

$$\delta W = \mathbf{E} \cdot d\mathbf{D} = \mathbf{E} \cdot \frac{d\mathbf{D}}{dt} dt$$

We obtain the work done over one period T of the electric field by integration. The expression of δW shows that the work per unit volume done over one period corresponds to the area enclosed by the curve $D(E)$. Writing the result as $W = p_a T$ yields the average power per unit volume, p_a , delivered to the dielectric medium. The power per unit volume received by the dielectric is

$$p(t) = \mathbf{E} \cdot \frac{d\mathbf{D}}{dt},$$

leading to the average power density:

$$p_a = \frac{1}{T} \int_0^T \epsilon_0 \epsilon_r^{(0)} \mathcal{E}_0^2 \omega (-\sin(\omega t - \delta) \cos \omega t) dt = \frac{1}{2} \epsilon_0 \epsilon_r^{(0)} \mathcal{E}_0^2 \omega \sin \delta$$

In order to obtain the same result using complex notations rather than real fields, we have to use the expression for the complex power density:

$$\mathcal{P} = \frac{1}{2} \mathbf{E} \cdot \left(\frac{d\mathbf{D}}{dt} \right)^*$$

where the star denotes the complex conjugate and the factor 1/2 corresponds to the time average of the sinusoidal field over one period.

Performing the calculation of the complex power density leads to $(d\mathbf{D}/dt)^* = (-i\omega\mathbf{D})^* = i\omega\mathbf{D}^*$ and $\mathcal{P} = \frac{1}{2} \mathcal{E}_0 i\omega \epsilon_r^{(0)} e^{-i\delta} \mathcal{E}_0^*$. The average power density delivered to the dielectric medium is then the real part of the complex power density:

$$p_a = \Re(\mathcal{P}) = \frac{1}{2} \epsilon_0 E_0^2 \epsilon_r^{(0)} \omega \sin \delta = \frac{1}{2} \epsilon_0 E_0^2 \omega \epsilon_r''$$

In conclusion, the imaginary part of the dielectric constant of a dielectric is linked to the power absorbed by the medium. After each cycle, the dielectric medium comes back to its initial electric state. The average power delivered to the dielectric leads to either an increase of its temperature if it is thermally isolated, or to heat radiation for an isothermal dielectric medium. Dielectric heating = *electronic heating* = *radio frequency heating* = *high-frequency heating Application*: Heating of the dielectric medium is the physical process used in heating by dielectric losses. The process is efficient if the frequency is chosen close to the resonance frequency, that is, the best operation conditions correspond to a frequency close to the frequency for which $\epsilon_r''(\omega)$ is maximum. This corresponds to the microwave frequency range $5 \text{ MHz} \leq \omega \leq 40 \text{ MHz}$, and electric fields up to 250 kV/m.

Heating with the microwave oven is based on this process:

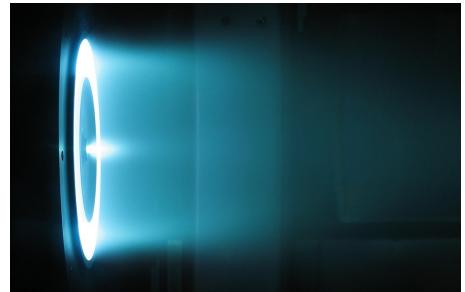
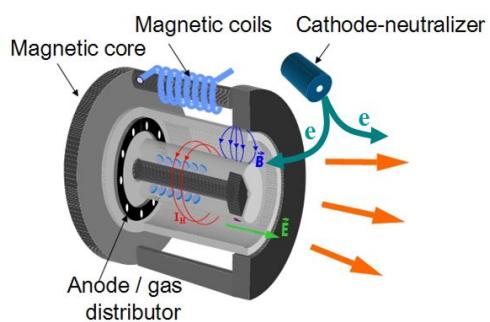
A microwave oven cooks food because the water molecules inside it absorb the microwave radiation and thereby heat up and heat the surrounding food. The electric field delivered by the microwave oven causes water molecules to vibrate, which generates intermolecular friction between the molecules of the food. The increased friction between the molecules results in heat. Microwave radiation will similarly heat up skin and other body parts. The radiation is harmful mostly to the parts of the body that cannot conduct the heat away very effectively—the eyes especially. If they were able to escape from the microwave oven, microwaves could affect your tissue. Modern microwave ovens are designed to allow essentially no leakage of microwaves, however. For instance, a metallic grid covers the transparent door window, acting as a Faraday cage preventing radiation leakage. The only time for concern would be if the door is broken or damaged, in which case the oven should not be used.

5.4 Summary and essential formulas

- Local field: $\mathbf{E}_l = \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0}$
- Polarizability α : defined from the dipole moment $\mathbf{p} = \epsilon_0 \alpha \mathbf{E}_l$
- Order of magnitude for α : volume of the atom/molecule
- Polarization mechanisms:
 - electronic $\mathbf{p} = \epsilon_0 \alpha \mathbf{E}_l$, with $\alpha = 4\pi R^3$
 - ionic: $\alpha = \frac{e^2}{\epsilon_0 Y |r_i^+ - r_i^-|}$
 - orientation: $\alpha = \frac{p_0^2}{3\epsilon_0 kT}$ for weak fields
- Clausius-Mossotti's relation: $\frac{n\alpha}{3} = \frac{\epsilon_r - 1}{\epsilon_r + 2}$
- Polarization: $\mathbf{P} = -n\mathbf{er}$
- Method to derive the dielectric constant: $\epsilon_r(\omega) = \epsilon'_r(\omega) + i\epsilon''_r(\omega)$ (or $\chi(\omega)$):
 - (1) Newton's equation for the electron
 - (2) Use a complex sinusoidal field of frequency ω
 - (3) Infer the complex susceptibility $\chi(\omega)$ from $\mathbf{P}(\omega) = \epsilon_0 \chi(\omega) \mathbf{E}(\omega)$.
 - (4) Infer/plot $\epsilon'_r(\omega) + i\epsilon''_r(\omega) = 1 + \chi(\omega)$ (look for resonance in the absorption $\epsilon''_r(\omega)$)
- Heating dielectrics: Complex power: $\mathcal{P} = \frac{1}{2} \mathbf{E} \cdot \left(\frac{d\mathbf{D}}{dt} \right)^*$
- Average power: $p_a = \frac{1}{2} \epsilon_0 E_0^2 \omega \epsilon''_r$

Chapter 6

Metal conductivity, conductors and electric currents, Hall effect



Introduction

As discussed in chapter 3, matter may be roughly classified into two families: conductors and insulators. Chapters 4 and 5 were devoted to the study of insulators. This chapter deals with the conductor materials in the presence of an electric field imposed by an external source that keeps two points of the conductor at different potentials. Thus, the electric field is present inside the conductor and accelerates the free electrons, which generates an electric current.

A stationary electric current develops, according to what is known as Ohm's law. This case is one of the subjects of the present chapter.

After reminders on conductors in static equilibrium, this chapter deals with transport phenomena that arise in conductors whenever the equilibrium is broken. In this aim, the property of metal conductivity is presented and modeled, both from the microscopic point of view and from the macroscopic point of view based on the Drude model. The metal conductivity is shown to be an important property of conductors, allowing us to write a constitutive relation of the material linking not only electric current and field (Ohm's law), but also macroscopic properties of the conductor, such as its resistance and its capacity.

In the last part of this chapter, we show how an external magnetic field acts on a conductor

that carries a current. The origin of the famous Laplace force resulting from the action of the magnetic field on a conductor carrying a current is discussed and shown to be due to the action of an electric field called the Hall field, acting on the ions of the metal. This field and the associated Hall effect are used for numerous applications, some of which are briefly presented.

6.1 Reminders about conductors in static equilibrium

The properties of conductors in electrostatic equilibrium were established in the PHY104 course:

- The electric field in a conductor at equilibrium is zero at every point inside the conductor. If Ω represents the volume of a conductive material, then

$$\mathbf{E}(\mathbf{x}) = \mathbf{0} \quad \forall \mathbf{x} \in \Omega$$

- The electrostatic potential is uniform inside a conductor in equilibrium.
- All the charge of a conductor in equilibrium is distributed on its surface $\partial\Omega$. The conductor is thus completely characterized by its surface charge density σ in $\partial\Omega$.
- The volume charge density is zero at every point inside the conductor

$$\rho(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \Omega$$

- The electric field is normal to the surface of a conductor in equilibrium. The surface of a conductor is an equipotential

$$\mathbf{E}(\mathbf{x}) = E(\mathbf{x})\mathbf{n}(\mathbf{x}) \quad \forall \mathbf{x} \in \partial\Omega .$$

- Coulomb's theorem states the relationship between the normal electric field to the surface of a conductor in equilibrium and the surface charge density $\sigma(\mathbf{x})$ it carries:

$$\mathbf{E}(\mathbf{x}) = \frac{\sigma(\mathbf{x})}{\epsilon_0} \mathbf{n}, \quad \mathbf{x} \in \partial\Omega. \tag{6.1}$$

Note that the notation σ for the surface charge density is also commonly used for a different physical quantity introduced below: the conductivity.

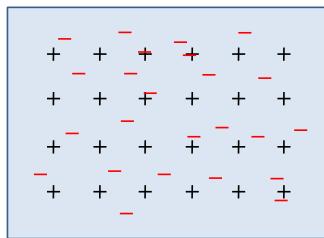
6.2 Microscopic description of conductivity

A proper microscopic description of the conductivity of a conductor requires arguments from statistical physics and quantum physics. However, it is possible to understand the origin of the conductivity from *phenomenological* arguments of classical physics, as carried out in the Drude model. This section reviews Drude's derivation.

6.2.1 Fixed and mobile charges

The role of a model for metals is to help understanding their properties and in particular their conductivity, which is due to the motion of electric charges.

A metal, for example copper, is modeled as a lattice of positively charged ionized atoms. These ions only move around their average position in the lattice; this is a small amplitude motion, that is, ions remain localized in the vicinity of their average position and can be assumed as fixed. Their volume charge density ρ_f is constant.



Each atom of the metal liberated at least one electron (Assume one electron per ion for the rest of the chapter). Electrons are mobile in the entire volume of the metal. They are called *charge carriers*. Their charge density is ρ_m . Both ρ_m and ρ_f are averaged densities. Below, we clarify the concept of averaging.

6.2.2 Average density of carriers

A volume element d^3x around point M contains at time t a number of charges

$$dN = n(\mathbf{x}, t)d^3x,$$

where $n(\mathbf{x}, t)$ denotes the volume density of charge carriers (in m^{-3}), not to be confused with a volume charge density (in C m^{-3}), at M and at time t . The density of charge carriers undergoes fluctuations in time due to thermal agitation. At time t , we obtain an average value of the density over a duration Δt

$$\bar{n}(\mathbf{x}, t) = \frac{1}{\Delta t} \int_{t-\Delta t/2}^{t+\Delta t/2} n(\mathbf{x}, t') dt'.$$

In a metal where no new charge carrier is generated, the average density $\bar{n}(\mathbf{x}, t)$ does not depend on M or t : $\bar{n}(\mathbf{x}, t) = n$. Thus, there is a link between the average density n and the charge density ρ_m of mobile charges and ρ_f of fixed charges:

$$\rho_m = -\rho_f = -ne,$$

where $-e$ is the negative charge of an electron and n denotes the volume density of charge carriers.

6.2.3 Average electric field

At point M , the microscopic electric field $\mathbf{e}(\mathbf{x}, t)$, defined as the electric field generated by all microscopic charges, undergoes fluctuations too because of thermal agitation of microscopic particles. These are very fast due to the motion of charge carriers. We define a temporally averaged electric field $\bar{\mathbf{e}}$ in the same way as the temporally averaged charge density:

$$\bar{\mathbf{e}}(\mathbf{x}, t) = \frac{1}{\Delta t} \int_{t-\Delta t/2}^{t+\Delta t/2} \mathbf{e}(\mathbf{x}, t') dt'.$$

This average electric field can be nonuniform (it depends on the coordinates of M) and time dependent. In the permanent regime, it is time-independent.

6.2.4 Spatially averaged electric fields

As seen in chapter 3, the fields that appear in Maxwell's equations as well as the sources (electric and magnetic fields, charge and current densities) are spatially averaged over a mesoscopic volume Ω_{ave} that scales as n^{-1} , the inverse of the density.

Spatially average field defined as

$$\mathbf{E}(\mathbf{x}, t) = \langle \mathbf{e}(\mathbf{x}, t) \rangle = \frac{1}{\Omega_{ave}} \iiint_{\Omega_{ave}} \mathbf{e}(\mathbf{x} + \mathbf{x}', t) d^3x'$$

vary slowly on the scale of thousands of particles ($n^{-1/3}$).

How does the spatially averaged electric field $\langle \mathbf{e}(\mathbf{x}, t) \rangle$ relate to the temporally averaged field $\bar{\mathbf{e}}(\mathbf{x}, t)$? In fact, all macroscopic quantities that appear in Maxwell's equations are averaged spatially *and* temporally, which smooths temporal fluctuations due to thermal agitation and spatial variations at the atomic scale.

The purpose of the time-averaging procedure is to eliminate the fast fluctuations of the physical quantities that are due to thermal agitation, that is, to the microscopic interaction between particles. This raises a conceptual difficulty that is related to the possibility of a macroscopic time-evolution of the fields, that must not be averaged out by the time-averaging procedure. Statistical physics postulates that physical quantities characterizing a macroscopic system can

be obtained by an ensemble average, where a large number of replicas of the system under investigation are prepared in the same conditions. The replicas are identical at the macroscopic scale but not at the microscopic scale as one cannot control the state of each individual particle of the system. Each replica will then undergo a different microscopic evolution due to the fluctuations associated with thermal agitation. However, the macroscopic evolution of all replicas is expected to be the same and the macroscopic physical quantities can then be obtained by performing an ensemble average, eliminating the fast fluctuations of the system.

Statistical physics relies on the *ergodicity principle* that states that a time-averaging procedure is equivalent to an ensemble average. This principle is a postulate, validated by experiments.

For electrodynamics, this means that the macroscopic fields in Maxwell's equations are given by the Lorentz (spatial) averaging procedure and that we accept that the resulting field equivalently represents the field from which the fluctuations due to thermal agitation are temporally averaged out:

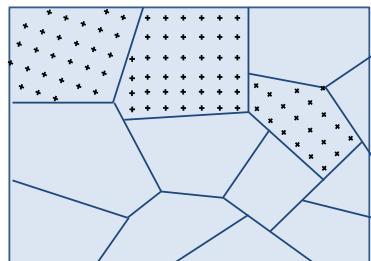
$$\mathbf{E}(\mathbf{x}, t) = \langle \mathbf{e}(\mathbf{x}, t) \rangle = \bar{\mathbf{e}}(\mathbf{x}, t).$$

In practice, we can view the macroscopic fields as spatially averaged quantities and the electrons in a metal as a fluid (or a gas) of electric charges. With this model in mind, we will see how to determine the conductivity of metals.

6.2.5 Collisions in a real metal

In a metal, a charge carrier may undergo collisions of various types. For instance, it can encounter an ion of the lattice or another charge carrier, but this is not the only possibility.

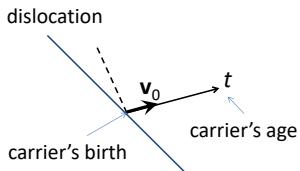
In a monocrystal, all the volume of the crystal forms a single perfectly regular crystal. A real metal is constituted by a juxtaposition of monocrystalline domains separated by dislocations of the lattice. A dislocation is simply a separation surface between two monocrystalline domains across which there is a discontinuity of the scalar potential. This discontinuity means that charge carriers will be accelerated through dislocations; in other words, their motion in the metal will undergo perturbations that can be described as collisions of electrons on the dislocations of the lattice.



The lattice also undergoes vibrations which can be described as waves, called *phonons*, similar to acoustic or optical waves, depending on their frequency. An electron will interact with these waves and will be scattered off the wave in a process that is analogous to a collision between two particles. Collisional processes are dominated by the type of collisions occurring the most frequently, which depends on the temperature.

6.2.6 Carrier lifetime

Electrons in a crystal, similarly to the case of molecules in a gas, are continuously subjected to collisions that are constantly changing their energy as well as the orientation of their velocities. As will be seen below, these microscopic processes are responsible for the Joule effect, which reflects the fact that electrons give away part of their kinetic energy under the form of heat (lattice vibrations).



Let us assume an electron undergoes a collision of any type and let us define the instant of the collision as the birth date of the electron. While the electron does not undergo another collision, the time t after the collision is defined as the age of the electron. Let us suppose that for an electron born at $t = 0$, the probability for a collision to occur during the time interval dt between times t and $t + dt$ is dt/τ , where τ is independent of the carrier's age t . It is also independent of the velocity that the electron had before the collision (all memory of its velocity prior to the collision is therefore lost) or after the collision (It is independent of v_0). Let us calculate the probability $p(t)$ that an electron reaches age t , that is, no collision occurs between 0 and t . In this aim, we will link $p(t+dt)$ to $p(t)$. If No collisions occurred between 0 and $t+dt$, it means that no collisions occurred between 0 and t , and No collisions occurred between the remaining time interval dt , so $p(t+dt)$ is equal to the product of probabilities that no collision occurs

$$\underbrace{p(t+dt)}_{\text{No coll. btw. 0 and } t+dt} = \underbrace{p(t)}_{0 \text{ and } t} \underbrace{(1 - dt/\tau)}_{t \text{ and } t+dt} \Rightarrow \frac{dp}{p} = -\frac{dt}{\tau}$$

After simple algebra, we obtained a well known ordinary differential equation for $p(t)$ whose solution is

$$p(t) = p_0 e^{-t/\tau}$$

where p_0 is determined from the normalization constraint

$$\int_0^{+\infty} p(t) dt = 1 \Rightarrow p_0 = 1/\tau$$

The age of the electron will lie between t and $t + dt$ with probability $p(t)$. Hence, we can calculate the average carrier's lifetime

$$\langle t \rangle = \frac{\int_0^{\infty} t p(t) dt}{\int_0^{\infty} p(t) dt} = \int_0^{\infty} \frac{t}{\tau} e^{-t/\tau} dt = \tau$$

We find an average lifetime equal to the constant τ introduced earlier.

$$\langle t \rangle = \tau = \Lambda / \bar{v} \quad (\text{mean free path / rms velocity})$$

We can thus interpret this time constant as the time between two collisions, that is the ratio of the mean free path by the rms velocity.

6.2.7 Motion of carriers

As seen in chapter 3, the Drude model, which allows for a qualitative description of the conductivity of metals considers electrons as a gas of classical particles.

In the absence of an electric field, the velocity \mathbf{v}_i of the i^{th} electron is randomly distributed and when averaged spatially, we obtain $\langle \mathbf{v}_i \rangle = \mathbf{0}$.

Under the influence of an electric field \mathbf{E} applied at $t = 0$, the motion of carriers is governed by Newton's second law:

$$m_e \frac{d\mathbf{v}_i}{dt} = -e\mathbf{E} \quad (6.2)$$

\mathbf{v}_i denotes the velocity of an electron of mass m_e and charge $-e$, that will be accelerated in the direction of the electric field according to the law:

$$\mathbf{v}_i(t) = \mathbf{v}_i(0) - \frac{et}{m_e} \mathbf{E}$$

Electrons cannot accelerate indefinitely. Of course, the Newton equation above, applied to a single electron, is only valid between two collisions, since an electron will give away part of its kinetic energy during each collision. The Newton equation must then be modified to take into account the fact that sooner or later, a collision will occur on a dislocation. In the previous section, we have seen that the probability for an electron of age t , with velocity $\mathbf{v}_i(t)$ to undergo a collision in the time interval dt between t and $t + dt$ is dt/τ . In order to take the effect of collisions into account, let us average the electron velocities (equivalent to performing a spatial average), and divide the electrons into two families:

- The fraction dt/τ of electrons that suffered a collision between t and $t + dt$. The contribution of these electrons to the average velocity at $t + dt$ is zero.
- The fraction $(1 - dt/\tau)$ of electrons that have not participated in any collision between t and $t + dt$. Their velocity has been increased by the electric field $\mathbf{v}_i(t + dt) = \mathbf{v}_i(t) - e\mathbf{E}dt/m_e$. Their contribution to the average velocity is therefore

$$\langle \mathbf{v}_i(t + dt) \rangle = \left(1 - \frac{dt}{\tau}\right) \left(\langle \mathbf{v}_i(t) \rangle - \frac{e\mathbf{E}dt}{m_e}\right)$$

which can be rewritten at first order in dt as

$$\langle \mathbf{v}_i(t+dt) \rangle - \langle \mathbf{v}_i(t) \rangle = \left(-\frac{e}{m_e} \mathbf{E} - \frac{1}{\tau} \langle \mathbf{v}_i(t) \rangle \right) dt. \quad (6.3)$$

In the limit $dt \rightarrow 0$ we recognize Newton's equation for the velocity of the gas of electrons $\mathbf{v}(t) = \langle \mathbf{v}_i(t) \rangle$:

$$m_e \frac{d\mathbf{v}}{dt} = -e\mathbf{E} - m_e \frac{\mathbf{v}}{\tau}, \quad (6.4)$$

where a friction term proportional to the average velocity \mathbf{v} of the electron gas, with a characteristic time τ , describes the effect of collisions, that is, the interaction between each electron and the rest of the crystal. This friction term in the equation of motion constitutes a phenomenological force describing the dissipation of energy under the form of heat (this is called the Joule effect).

Equation (6.4) admits a steady state solution for the velocity

$$\mathbf{v} = -\frac{e\tau}{m_e} \mathbf{E} = -\mu_e \mathbf{E}, \quad (6.5)$$

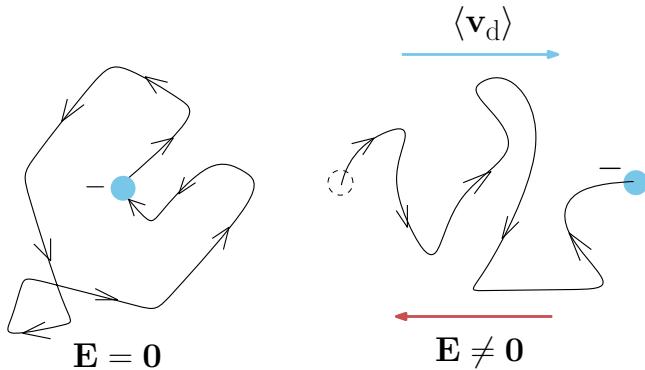
where $\mu_e = e\tau_e/m_e$ is the electron **mobility**. As seen in chapter 3, for a given material, this quantity determines the proportionality constant between the external electric field and the steady state average velocity of the electrons. It is this average velocity that supports the steady current developing in a conductor whenever an electric field is applied.

Before describing macroscopic properties of this current, note the connection between the average time between collisions, introduced as a microscopic feature of the metal, and the trend of the macroscopic electron gas to return to equilibrium: If we suppose that, starting from a steady state, the electric field is turned off at $t = 0$, the averaged velocity will then decay exponentially to zero according to $\mathbf{v}(t) = \mathbf{v}_0 \exp(-t/\tau)$, with a typical characteristic time τ . The latter quantity then represents the time it takes for the system to come back to the equilibrium, and it is therefore called relaxation time. The relaxation time is equal to the average time between two collisions, in keeping with the intuitive idea that the only way for the electron gas to reach equilibrium is by means of collisions.

6.3 Electric currents

We are now in a position to clarify what is meant by saying that in a conductor at equilibrium, the electrons are static inside the conductor and do not move. In fact, this is only true in average, since locally, each electron will feel the electric fields generated by the nuclei, by the other electrons, by charged impurities, etc. The electrons are not at rest, but constantly jiggling around due to all these interactions. After each collision, the electron momentum acquires a new, random direction so that in the absence of an external electric field, the electron describes a random walk, similar to what happens to a molecule in a perfect gas. In equilibrium, the average velocity of an electron in a conductor is zero, $\langle \mathbf{v} \rangle = \mathbf{0}$.

Despite its very large velocity between collisions, the electron does not advance in any particular direction. To have an idea of how fast electrons move in metals, recall that for a classical gas the averaged kinetic energy is $1/2m_e\langle \mathbf{v}^2 \rangle = 3/2k_B T$, where k_B is the Boltzmann constant and T the temperature. At room temperature ($T = 292$ K), we have $\sqrt{\langle \mathbf{v}^2 \rangle} \sim 1 \times 10^5$ m s $^{-1}$. This value is however an underestimation, since electrons in metals behave according to quantum laws. In copper, for example, the velocity between collisions reach 1×10^6 m s $^{-1}$ due to Pauli's exclusion principle.



Let us now consider an out-of-equilibrium situation, in which an external electric field \mathbf{E} is present inside a conductor. Due to this field, between two collisions the electron will be accelerated in the direction opposite to the electric field. After multiple collisions, the electron has drifted in this direction so that the electron acquires a non-zero average velocity $\langle \mathbf{v}_d \rangle$, called drift velocity, given by Eq. (6.5). In general, it is orders of magnitude smaller than the thermal velocity between collisions, but it is this velocity, always in the same direction, that is responsible for the appearance of electrical currents.

Averaging charge velocities can be performed not only in a conductor but in any medium, for instance a plasma, that contains both positive and negative free charges. This allows us to define the current density that appears in Maxwell's equations.

Definition 6.1: Conduction current density

Let n_+ (resp. n_-) be the density of positive (resp. negative) charges, each of charge q_+ , (resp q_-) in a medium and \mathbf{v}_+ (resp. \mathbf{v}_-) the average drift velocity of these charges. The conduction current density is defined as

$$\mathbf{J} = q_+ n_+ \mathbf{v}_+ + q_- n_- \mathbf{v}_- \quad (6.6)$$

and is measured in C s $^{-1}$ m $^{-2}$.

Remark

For an electrically neutral medium such as a conductor or a plasma, the total charge density is generally zero, which is expressed as

$$n_+ q_+ + n_- q_- = 0.$$

For a conductor where the mobile charges are electrons moving in a lattice of fixed ions of opposite charge, we have $q_- = -e$ and $q_+ = e$, hence, the neutrality condition implies that the density of ions and electrons are equal

$$n_+ = n_- = n$$

. However, a current exists when the velocity \mathbf{v}_+ of positive charges, usually much heavier than electrons, can be neglected compared to the velocity of electrons $\mathbf{v}_- = \mathbf{v}$. This is the case in conductors and the conduction current density is due to the *free* electrons,

$$\text{vect} J = -ne\mathbf{v}.$$

Recall that its flux through a surface defines the electric current $I = \Phi_{S,\mathbf{J}}$ (See chapter 1). Also recall from chapter 1 that the current density satisfies the charge conservation equation,

$$\boxed{\nabla \cdot \mathbf{J}(\mathbf{x}, t) + \frac{\partial \rho(\mathbf{x}, t)}{\partial t} = 0}, \quad (6.7)$$

where $\rho(\mathbf{x}, t)$ denotes the total charge density in the medium. From Eq. (6.7), we see that a current density compatible with a steady state regime ($\partial \rho / \partial t = 0$) is such that

$$\nabla \cdot \mathbf{J}(\mathbf{x}) = 0, \quad (6.8)$$

which means that the total current across a closed surface must be zero in the steady state. A consequence of this statement is the conservation of the electric current through the section of a current tube. For instance, a cylindrical wire constitutes a current tube as the electron velocity, and so the current density \mathbf{J} are parallel to the lateral section of the wire (otherwise, electrons would escape). The flux of \mathbf{J} through any section, that is the electric current through any section, is preserved¹

6.4 Ohm's law - Microscopic form

According to the Drude model, the current density \mathbf{J} in a conductor having a density n of free electrons per unit volume may be written as $\mathbf{J} = -ne\langle \mathbf{v}_i \rangle$. Using Eq. (6.5), we see that the steady state current density is proportional to the electric field \mathbf{E} :

$$\boxed{\mathbf{J} = ne\mu_e \mathbf{E} = \frac{ne^2\tau_e}{m_e} \mathbf{E} = \sigma \mathbf{E}.} \quad (6.9)$$

¹Indeed, the difference between the exit- and the entrance-sections represent the flux of \mathbf{J} through the closed surface {entrance + lateral + exit surfaces }, the flux through the lateral surface being zero. Ostragradski's theorem tells us that this flux is the volume integral of $\nabla \cdot \mathbf{J}$, hence it is equal to zero.

Equation (6.9) is the microscopic form of **Ohm's law**. The quantity

$$\sigma = \frac{ne^2\tau}{m}$$

is the **electrical conductivity** of the material, whose units in the S.I. system is the Siemens per meter (Sm^{-1}), where a Siemens corresponds to $1\text{S} = 1\text{C m}^{-1}\text{s}^{-1}\text{V}^{-1}$. The resistivity ϱ is defined as the inverse of the conductivity, $\varrho = 1/\sigma$ (this notation should be used carefully since it is almost the same as that used for the charge density). A good insulator (glass, plastics, dry air) has a conductivity of the order of $\sigma = 10^{-12} - 10^{-16}\text{ Sm}^{-1}$, whereas good conductors (gold, copper, silver) have conductivities around $\sigma = 1 \times 10^8\text{ Sm}^{-1}$. The electrical conductivity is one of the physical properties having one of the largest variations in nature (24 orders of magnitude may be spanned between a good conductor and a good insulator).

Remark

Note that the steady state current density must satisfy $\nabla \cdot \mathbf{J} = 0$ and, according to Ohm's law, $\sigma \nabla \cdot \mathbf{E} = 0$ (if the conductivity is homogeneous). From Gauss's law, this implies $\rho = 0$ and so the charge density inside a conductor must vanish at steady state. This property not only extends the known property for conductors at equilibrium, but also means that the electric potential in a metal satisfies Laplace's equation

$$\Delta V = 0.$$

Methods of potential theory to find the potential, the field $\mathbf{E} = -\nabla V$ and the current $\mathbf{J} = \sigma \mathbf{E}$ apply in a homogeneous ohmic conductor similarly to vacuum.

6.4.1 Ohm's law in the case of a time-dependent electric field

The microscopic Ohm law $\mathbf{J} = \sigma \mathbf{E}$ is valid at steady state. In the more general case of a time dependent electric field, we will see that the conductivity σ of the medium depends on how fast the field oscillates. In other words, the conductivity becomes frequency-dependent. In order to see this, suppose that the electric field has a sinusoidal time-dependence of the form

$$\mathbf{E}(t) = \mathbf{E}_0 \cos(\omega t)$$

and let us find a solution for the equation of motion of an electron accelerated by this electric field

$$\frac{d\mathbf{v}}{dt} = -\frac{e}{m_e} \mathbf{E}_0 \cos(\omega t) - \frac{\mathbf{v}}{\tau}.$$

To solve this type of equation, it is convenient to write the electric field and the electron velocity as

$$\mathbf{E} = \text{Re}\{\underline{\mathbf{E}}\}, \quad \mathbf{v} = \text{Re}\{\underline{\mathbf{v}}\},$$

where the complex electric field and velocity are given by $\underline{\mathbf{E}} = \mathbf{E}_0 e^{-i\omega t}$ and $\underline{\mathbf{v}} = \mathbf{v}_0 e^{-i\omega t}$, respectively. Since the equation of motion is linear, we can write

$$\frac{d\underline{\mathbf{v}}}{dt} = -\frac{e}{m_e} \underline{\mathbf{E}} - \frac{\underline{\mathbf{v}}}{\tau}$$

so that the real part of the latter equation (on both sides) gives the original equation of motion. We can therefore solve the equation for the complex velocity, which is particularly simple since $\frac{d\underline{\mathbf{v}}}{dt} = -i\omega \mathbf{v}_0 e^{-i\omega t}$ so that

$$i\omega \mathbf{v}_0 e^{-i\omega t} = -\frac{e}{m_e} \mathbf{E}_0 e^{-i\omega t} - \frac{\mathbf{v}_0 e^{-i\omega t}}{\tau}.$$

The complex exponentials cancel out on both sides and we obtain for the complex amplitude of the velocity

$$\mathbf{v}_0 = -\frac{e\tau/m_e}{(1-i\omega\tau)} \mathbf{E}_0.$$

Since the complex current density is given by $\underline{\mathbf{J}} = -en\underline{\mathbf{v}}$, we get the generalized Ohm law for a sinusoidal regime:

$$\underline{\mathbf{J}} = \underline{\sigma}(\omega) \underline{\mathbf{E}}, \quad (6.10)$$

where the complex conductivity is given by

$$\underline{\sigma}(\omega) = \frac{\sigma_0}{1-i\omega\tau} = \frac{ne^2\tau/m_e}{1-i\omega\tau}, \quad (6.11)$$

where $\sigma_0 = ne^2\tau/m_e$ is the static conductivity. From this result, the real current density writes

$$\mathbf{J} = \operatorname{Re}\{\underline{\mathbf{J}}\} = \operatorname{Re}\left\{\frac{ne^2\tau/m_e}{(1-i\omega\tau)} \mathbf{E}_0 e^{-i\omega t}\right\} = \frac{\sigma_0}{\sqrt{1+(\omega\tau)^2}} \mathbf{E}_0 \cos(\omega t - \tan^{-1}(\omega\tau)).$$

We see that if $\omega \ll \frac{1}{\tau}$, we can write a static Ohm's law at each time, $\mathbf{J} \approx \sigma_0 \mathbf{E}_0 \cos(\omega t) = \sigma_0 \mathbf{E}(t)$.

6.4.2 About the validity of the Drude model

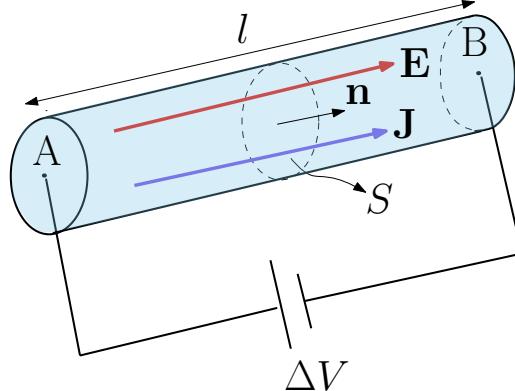
The Drude model is a simple way to describe conductivity and Ohm's law in terms of collisions between classical particles. However a more complete, quantitative description of conductivity requires us to take into account the quantum behavior of electrons in solids. At a microscopic level, the notion of a well-defined trajectory for the electrons must be abandoned, and replaced by probability waves propagating through the crystal. Indeed, the classic model of Drude fails in a number of points discussed below.

- The Drude model fails to predict the correct thermal conductivity κ of a metal and its relationship with the electrical conductivity σ . Interestingly, the model also fails to predict the correct kinetic energy of the electron gas. These two errors almost compensate each other so that the ratio κ/σ is, accidentally, only off by a factor of two from the correct value $\kappa/\sigma = \frac{\pi^2}{3}[k_B/e]^2T$ (Wiedemann-Franz law) which is well explained by quantum mechanics.
- In reality, not all electrons can participate in the conductivity. Electrons with small kinetic energies are excluded from any collision mechanism due to Pauli's exclusion principle, and for the same reasons these electrons may not accelerate under the presence of an electric field. With a complete quantum treatment, one may justify a transport equation similar to Drude's Eq. (6.5), but in which τ is not anymore the same for every electron but instead an average over the electron statistical distribution, which differs from the Maxwell-Boltzmann distribution used for a classical gas.
- Another key issue is related to the electron mean free path (the average distance traveled between two collisions) in a solid. Naively, one may think that the trajectory of an electron will be modified each time it encounters an ion from the crystal lattice. However, the experimental values for the time between collisions are such that the mean free path is much larger than the distance between atoms in the crystal, and this cannot be explained with a classical picture. In reality, electrons behave like waves which are not perturbed by the presence of atoms of the lattice. Instead, these waves are only perturbed by deviations from a perfect crystal (atomic vibrations, impurities, etc.).
- Finally, there is a whole class of materials relevant for modern technologies called semiconductors, for which the conductivity is intermediate between that of an insulator and that of a conductor, and that can be tuned by several orders of magnitude by small perturbations (light excitation, electric fields, incorporation of minute proportions of chemical impurities). This versatility is at the heart of the major technological revolution of the 20th century (transistors, diodes, LEDs, solar cells and semiconductor lasers are a few examples of devices based on semiconducting materials). In solids in general, an electron does not have an arbitrary energy, but instead only energies that fall within bounded intervals (called bands) separated by energy gaps where no possible states are allowed for the electrons. In a semiconductor, a gap exists between non-conductive and conductive states which is of the order of 1 eV. This explains why, for example, increasing the temperature (i.e. the available thermal energy) increases the conductivity, which is the opposite to what is observed in metals. The existence of this energy gap can only be explained with quantum mechanics.

6.5 Ohm's law - Macroscopic form

Let us now consider a wire made of a conductive material of length l , section S and conductivity σ . A difference of electric potential $\Delta V = V_A - V_B$ is imposed between the two ends of the

wire, as shown in the figure below.



This potential difference generates an electric field \mathbf{E} , and in consequence, a current density $\mathbf{J} = \sigma\mathbf{E}$. At steady state, $\nabla \cdot \mathbf{J} = 0$ so that the current entering the wire through A equals the output current at B . If the section S is constant, the current density \mathbf{J} as well as the electric field \mathbf{E} are uniform inside the conductor. Then

$$\Delta V = V_B - V_A = \int_A^B \mathbf{E}(\mathbf{x}) \cdot d\mathbf{x} = E l.$$

The current density therefore writes

$$\mathbf{J} = \sigma\mathbf{E} = \sigma \frac{\Delta V}{l} \mathbf{n},$$

where \mathbf{n} is the direction normal to the transverse section S and along the electric field. The current through the wire is then

$$I = \mathbf{J} \cdot \mathbf{n} S = \Delta V \frac{\sigma S}{l}$$

which leads to the macroscopic Ohm law.

Definition 6.2: Ohm's law (1827)

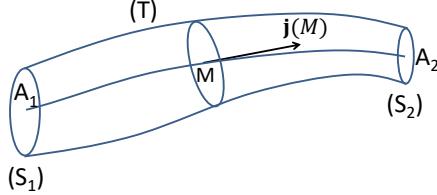
The potential difference between two extremities A and B of a conductor of section S , length l and conductivity σ is proportional to the current going through it:

$$\Delta V = \frac{l}{\sigma S} I = RI, \quad (6.12)$$

where $R = \frac{l}{\sigma S} = \frac{l\rho}{S}$ is the **resistance** of the conductor between A and B . In the S.I. system, the unit for the resistance is the Ohm (Ω), with $1\Omega = 1\text{VA}^{-1}$.

6.5.1 Relation between resistance and capacity

Consider a current tube in a conductor, that is, the lateral surface of the tube is tangent to the current density vector at each point of the surface. Two sections of the current tube, S_1 and S_2 are orthogonal to \mathbf{J} and limit the volume under consideration. Between the surfaces, we have a metallic medium of conductivity σ



In the steady state regime, due to Ohm's law $\mathbf{J} = \sigma\mathbf{E}$, the current is proportional to the electric field, hence, a current tube is also a field tube. The lateral surface of the tube is tangent to the electric field at each point of the surface. The surfaces S_1 and S_2 are then also orthogonal to the electric field, thus, they are equipotential surfaces, say S_1 at potential V_1 and S_2 at potential V_2 .

The resistance of this *tube* is defined as the ratio $R = (V_1 - V_2)/I$ where

$$\left. \begin{aligned} V_1 - V_2 &= \int_{A_1}^{A_2} \mathbf{E} \cdot d\mathbf{l} \\ I &= \iint_S \mathbf{J} \cdot \mathbf{n} dS, \quad \mathbf{J} = \sigma\mathbf{E} \end{aligned} \right\} R = \frac{\int_{A_1}^{A_2} \mathbf{E} \cdot d\mathbf{l}}{\iint_{S_1} \sigma\mathbf{E}_1 \cdot \mathbf{n}_1 dS_1}$$

Note that charge conservation, i.e., $\nabla \cdot \mathbf{J} = 0$ implies that the current I can be calculated as the flux of the current vector through any section of the field tube (here through S_1 ; it would yield the same result if calculated through S_2).

The capacity of the virtual capacitor made of the two surfaces S_1 at potential V_1 and S_2 at potential V_2 reads $C = Q/(V_1 - V_2)$ where Q denotes the free charge on surface S_1 , that is

$$Q = \iint_{S_1} \varsigma_1 dS_1$$

where the surface density of free charge² ς_1 is obtained from Coulomb's theorem,

$$\varsigma_1 = \epsilon_0 \mathbf{E}_1 \cdot \mathbf{n}_1.$$

Thus,

$$C^{-1} = \frac{V_1 - V_2}{Q} = \frac{\int_{A_1}^{A_2} \mathbf{E} \cdot d\mathbf{l}}{\iint_{S_1} \epsilon_0 \mathbf{E}_1 \cdot \mathbf{n}_1 dS_1}.$$

²not to be confused with the conductivity σ

From the resistance and the capacity of this virtual capacitor, we obtain

$$\sigma R = \frac{\epsilon_0}{C} \quad \text{i.e.} \quad RC = \rho \epsilon_0 = \frac{\epsilon_0}{\sigma} \quad (6.13)$$

The product of the resistance by the capacity of a conductor is equal to the product of the metal resistivity by the dielectric constant of vacuum.

Suggestions for exercises:

- Resistance between two concentric cylinders (Exercise textbook: Ex. 4.17 p. 320)
- Resistance between two concentric spheres (Exercise textbook: Ex. 4.18 p. 322)

6.5.2 Charge neutrality in a conductor

We have seen that inside a conductor in electrostatic equilibrium, the electric field must vanish, so that $\mathbf{E} = \mathbf{0}$ and the charge density ρ is zero as well at every point. We will show now that more generally, when a time-varying electric field is present, the charge neutrality on a conductor may or not be respected depending on the frequency of the electric field with respect to the characteristic timescales at which charge densities can oscillate in the metal. Let us first suppose a sinusoidal electric field of the form

$$\mathbf{E}(t) = \mathbf{E}_0 \cos(\omega t) = \operatorname{Re}\{\underline{\mathbf{E}}\}$$

for which the current density is given by Ohm's law in the sinusoidal regime (6.10)

$$\mathbf{J}(t) = \operatorname{Re}\{\underline{\mathbf{J}} \cdot \underline{\mathbf{E}}\} = \operatorname{Re}\left\{\frac{\sigma_0}{1 - i\omega\tau} \underline{\mathbf{E}}\right\}$$

Now, recall that the charge density ρ must satisfy the charge-conservation equation

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

and writing $\rho = \operatorname{Re}\{\underline{\rho}\} = \operatorname{Re}\{\rho_0 e^{-i\omega t}\}$, we may instead look for the equation that ρ_0 must satisfy

$$\frac{\partial \underline{\rho}(\mathbf{x}, t)}{\partial t} + \nabla \cdot \underbrace{\mathbf{J}(\mathbf{x}, t)}_{\underline{\sigma}(\omega)\underline{\mathbf{E}}} = 0$$

which writes

$$-i\omega \underline{\rho} + \underline{\sigma}(\omega) \nabla \cdot \underline{\mathbf{E}} = 0$$

and according to Gauss's law, $\nabla \cdot \underline{\mathbf{E}} = \underline{\rho}/\epsilon_0$

$$-i\omega \rho_0 + \frac{\underline{\sigma}(\omega)}{\epsilon_0} \rho_0 = -i\omega \rho_0 + \frac{\sigma_0}{(1 - i\omega\tau)\epsilon_0} \rho_0 = 0$$

or, equivalently,

$$\omega^2 \rho_0 + i \frac{\omega}{\tau} \rho_0 - \underbrace{\frac{\sigma_0}{\epsilon_0 \tau}}_{\omega_p^2} \rho_0 = 0,$$

where we define the plasma frequency as

$$\omega_p = \sqrt{\frac{\sigma_0}{\epsilon_0 \tau}} = \sqrt{\frac{ne^2}{\epsilon_0 m_e}}$$

and the latter equation rewrites

$$\left(\omega^2 - \omega_p^2 - i \frac{\omega}{\tau} \right) \rho_0 = 0.$$

Typically, in metals, we have $1/\tau \ll \omega_p$ so that $\omega = \omega_p$ is a solution of the above equation for which $\rho_0 \neq 0$. Thus, an electric field oscillating at the plasma frequency can break the charge neutrality of the conductor, maintaining charge waves that naturally oscillate in phase with the electric field.

In the general case $\omega \neq \omega_p$, we see that the charge density satisfies the equation of motion of a damped harmonic oscillator of eigenfrequency ω_p

$$\boxed{\frac{\partial^2 \rho}{\partial t^2} + \frac{1}{\tau} \frac{\partial \rho}{\partial t} + \omega_p^2 \rho = 0}. \quad (6.14)$$

At low frequencies ($\omega \ll \frac{1}{\tau} \ll \omega_p$), the first term may be neglected with respect to the other two and the equation writes

$$\frac{\partial \rho}{\partial t} + \underbrace{\tau \omega_p^2}_{=\sigma_0 \epsilon_0} \rho = 0.$$

In this regime, if a perturbation breaks the charge neutrality in a conductor at $t = 0$, the charge density will evolve as

$$\rho(\mathbf{x}, t) = \rho_0(\mathbf{x}) e^{-t/\tau_d}$$

so that the medium recovers its neutrality exponentially fast, as e^{-t/τ_d} . The characteristic decay time $\tau_d = \sigma_0 / \epsilon_0$ is called the dielectric relaxation time. Any charge imbalance excited by a low-frequency perturbation is therefore neutralized within a time τ_d . Note that from (6.13), the dielectric relaxation time can also be written as $\tau_d = RC$. In metals, this time can be extremely short, of the order of $\tau_d = 1.3 \times 10^{-18}$ s, whereas in an insulator such as quartz, we have $\tau_d \sim 1 \times 10^7$ s, which is of the order of one year.

6.6 Electric energy and power

6.6.1 Power supplied by a battery

Suppose that a battery imposes a potential difference $\Delta V = V_A - V_B$ between the two ends A and B of a conductor, thus generating a current $I = \frac{\Delta V}{R}$ where R is the total resistance between A and B . In order to keep the current going, the battery must be constantly supplying the energy required for a charge q to recover its potential energy drop after each turn. This corresponds to an electric power given by

$$P = \frac{dq}{dt} \Delta V = I \Delta V \quad (6.15)$$

6.6.2 Joule effect (1860) - Power dissipated in a resistance

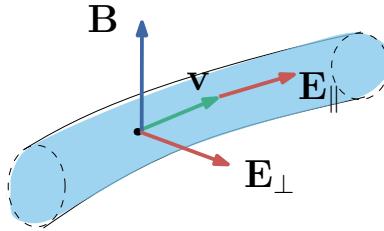
At steady state, the power supplied by the battery, which imposes a potential difference ΔV between the two ends of a conductor of resistance R , equals the amount of power dissipated in the resistance. Microscopically, this is due to the relaxation of the electron kinetic energy inside the crystal. During collisions, part of the energy is transferred to the vibrational modes of the atoms in the crystal, increasing its temperature. The electric energy thus generates heat inside the conductor. The power dissipated must be equal to the power supplied by the battery, so that $P = I \Delta V$, and according to Ohm's law (6.12), $\Delta V = RI$, and so

$$P = I \Delta V = RI^2 = \Delta V^2 / R. \quad (6.16)$$

We see that the dissipated energy is proportional to the square of the current. One may apply this principle to generate heat in an oven, to generate hot air or to heat the filament in a light bulb. For some applications, the Joule effect must be minimized in order to keep a good energetic efficiency. This is the case of a transmission line, where any loss of energy between a power plant and a city must be minimized. This is why transmission lines operate at very high voltages, thus minimizing the current for a given electrical power.

6.7 The Hall effect

The Hall effect is a consequence of the Lorentz force acting on charge carriers. It manifests itself as the appearance of a transverse electric field (perpendicular to the velocity of the electrons) whenever a current flows in the presence of an external magnetic field \mathbf{B} .



If \mathbf{v} is the average velocity of an electron in a conductor, we can write an equation of motion which is a generalization of Drude's model in order to include the contribution of the magnetic force, that is,

$$m \frac{d\mathbf{v}}{dt} + \frac{m}{\tau} \mathbf{v} = -e\mathbf{E} - e\mathbf{v} \times \mathbf{B}$$

and at steady state,

$$\mathbf{v} = -\frac{e\tau}{m}\mathbf{E} - \frac{e\tau}{m}\mathbf{v} \times \mathbf{B}.$$

We can write the electric field as $\mathbf{E} = \mathbf{E}_{\parallel} + \mathbf{E}_{\perp}$, where \mathbf{E}_{\parallel} is the electric field component parallel to the velocity of the electrons, and \mathbf{E}_{\perp} the transverse component. Projecting the equation of motion onto the parallel direction yields

$$\mathbf{v} = -\frac{e\tau}{m}\mathbf{E}_{\parallel} \rightarrow \mathbf{J} = \frac{ne^2\tau}{m}\mathbf{E}_{\parallel} = \sigma\mathbf{E}_{\parallel},$$

which is the usual Ohm's law, where n is the density of free electrons and σ is the conductivity. The projection along the direction transverse to \mathbf{v} imposes

$$\mathbf{0} = -\frac{e\tau}{m}\mathbf{E}_{\perp} - \frac{e\tau}{m}\mathbf{v} \times \mathbf{B}.$$

The transverse field \mathbf{E}_{\perp} is also called the Hall field and is given by

$$\mathbf{E}_H = -\mathbf{v} \times \mathbf{B}$$

(6.17)

The building up of this Hall field is required so that at steady state, the trajectory of the electron is not perturbed by the external magnetic field. It comes from an inhomogeneous accumulation of electrons at the surface of the conductor.

Collecting the parallel and the perpendicular components of the field, we obtain

$$\mathbf{E} = \mathbf{E}_{\parallel} + \mathbf{E}_{\perp} = \frac{\mathbf{J}}{\sigma} - \frac{\mathbf{J}}{-ne} \times \mathbf{B}$$

which takes the form

$$\mathbf{E} = \varrho\mathbf{J} - R_H\mathbf{J} \times \mathbf{B} \quad \text{with} \quad \varrho = \frac{1}{\sigma} \quad \text{and} \quad R_H = -\frac{1}{ne}$$

R_H is called the Hall coefficient.

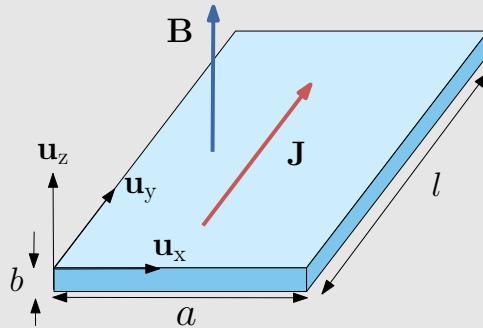
Historically, the Hall effect was of crucial importance for the understanding of the conduction properties of semiconducting materials, and is nowadays widely used as a means to measure magnetic fields.

Example 6.1 - The Hall effect

Consider a conductive plate of length l and rectangular section of width a and thickness b . The plate carries a uniform current density \mathbf{J} that is parallel to its length (along the y -axis). The conductor has n free charges (q) per unit volume.

- a.** Write the relationship between the current density \mathbf{J} and n .

Now a magnetic field \mathbf{B} is applied along the z -axis.



- b.** What is the effect of the field \mathbf{B} on free charges? What forces are they subjected to and what happens on the side walls $x = 0$ and $x = a$? On which wall is there an accumulation of charges q ?

The walls on which there is an accumulation or deficit of charges behave like the plates of a flat capacitor and an electric field is established between them \mathbf{E}_H (The Hall field).

- c.** What is the effect of the \mathbf{E}_H field on free charges?
- d.** For what value of the \mathbf{E}_H field is an equilibrium achieved in which the effect of the magnetic field \mathbf{B} is canceled? What is then the trajectory of free charges?
- e.** What is the Hall voltage V_H that appears between the walls in terms of the Hall constant $R_H = -1/nq$? How can the Hall effect be used to measure the value of an unknown magnetic field (Hall effect probe)?

Solution

- a.** The current density is given by $\mathbf{J} = nqv$, where \mathbf{v} is the speed of the free charges on the conductor. Since the current flows in the direction \mathbf{u}_y , we have $\mathbf{v} = v\mathbf{u}_y$ and then,

$$\mathbf{J} = qnv\mathbf{u}_y.$$

- b.** The magnetic field \mathbf{B} exerts an additional force on a charge, given by the magnetic force

$$\mathbf{F}_B = q\mathbf{v} \times \mathbf{B} = qv\mathbf{u}_y \times B\mathbf{u}_z = qvB\mathbf{u}_x.$$

Then, the magnetic force accelerates the charges along the axis \mathbf{u}_x perpendicular to the current. If $q > 0$, the trajectory of the charges will be diverted to the wall $x = a$, while if $q < 0$, the charges will be diverted to the wall $x = 0$. There is then an accumulation of charges on one of the walls and a deficit of charges on the opposite wall.

- c. The electric field $\mathbf{E}_H = -E_H \mathbf{u}_x$ between the plates generates an additional electrostatic force on the free charges, given by

$$\mathbf{F}_H = q\mathbf{E}_H = -qE_H \mathbf{u}_x.$$

Note that if $q > 0$, the plate $x = d$ will be positively charged and $E_H > 0$. If $q < 0$, $E_H < 0$. In all cases, the force associated with the Hall field will be in the direction of $-\mathbf{u}_x$, that is, in the direction opposite to the magnetic force.

- d. For the effect of the magnetic field on the charges to be null in the stationary regime, we must have

$$\mathbf{F}_B + \mathbf{F}_H = 0,$$

that is,

$$qvB\mathbf{u}_x - qE_H \mathbf{u}_x = 0 \rightarrow E_H = vB.$$

In the stationary regime, the charges move along a straight line parallel to the current density \mathbf{J} .

- e. The Hall voltage V_H that appears between the side walls is

$$V_H = V(x = 0) - V(x = a) = -aE_h = -avB.$$

Note that if I is the current flowing through the plate, then $I = abJ = abnqv$, so $v = I/(abnq)$ and

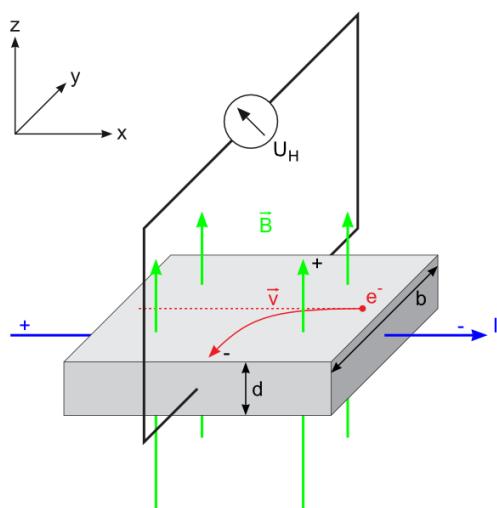
$$V_H = -\frac{BI}{bnq} = R_H \frac{BI}{b},$$

where $R_H = -1/(nq)$ denotes the Hall resistance, which depends solely on the density of free charges in the conductor. Thus, if it is desired to determine the magnitude of an unknown magnetic field \mathbf{B} , a current I is applied in the plate perpendicular to the magnetic field, the voltage V_H that appears between the side walls is measured and so,

$$B = \frac{V_H b}{IR_H}.$$

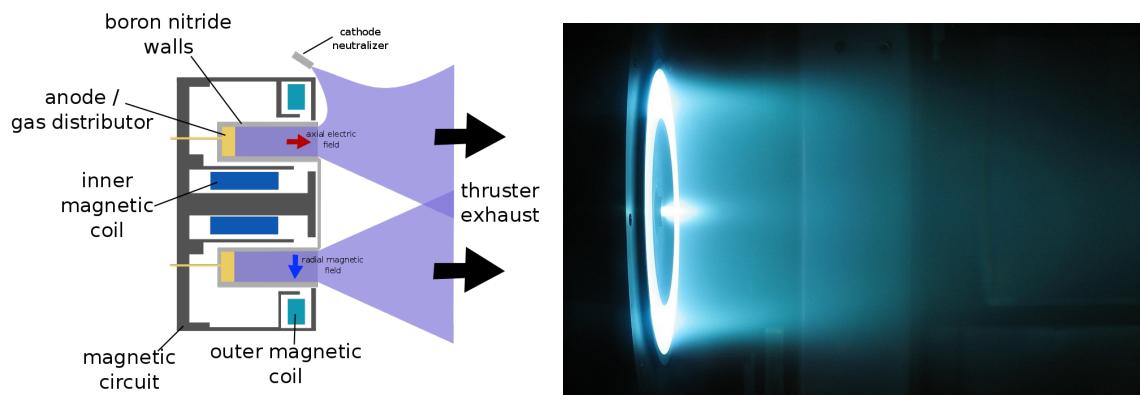
6.7.1 Applications of the Hall effect

The Hall effect is used in numerous devices of everyday life as well as in modern technology. It led to the development of Hall Effect probes which allow us to measure magnetic fields since the Hall effect provides a direct relationship between a current flowing through a conductor and the magnetic field applied to the conductor perpendicularly to the current. The Hall effect can be used to build compasses as it is sufficient to measure a voltage difference to infer a magnetic field magnitude. The Hall effect can also be used to build contactless switches to power certain objects, e.g., in the automotive industry for engine ignition.



Hall probes can be used for various purposes: Here is a non-exhaustive list of examples:

- Detection of rotation speed: can be used in bicycle wheels, as speedometers in the automotive world, electronic types of ignition systems, and control of gear teeth rotation.
- Detection of movement (smart phones, paintball guns, some GPS systems.)
- Position and motion sensors: mainly used in a DC motor, often of brushless type.
- In the automotive world: fuel injection and ignition. Wheel rotation sensors for anti-lock braking systems.
- For spacecraft propulsion: Hall-effect thruster are used to propel spacecrafts, after they get into orbit.



The principle of a Hall thruster for spacecraft propulsion is the following:

- Electrons are trapped in a radial magnetic field.
- The electrons ionize a propellant (Usually a xenon gas, which has heavy atoms and low ionization potential to be easily ionized).
- Ions are accelerated to produce thrust
- Ions are then neutralized in the plume.
- The exhaust velocity ranges within 10-80 km/s

6.8 Summary and essential formulas

- In a metal in the presence of an electric field (out of static equilibrium), electrons will acquire a non-zero average velocity \mathbf{v} , which according to the Drude model is given by

$$\mathbf{v} = \frac{-e\tau}{m_e} \mathbf{E} = -\mu_e \mathbf{E},$$

where $-e$ is the electron charge, τ the mean time between two collisions, and m_e the electron mass. This allows us to define the mobility $\mu_e = \frac{e\tau}{m_e} > 0$. In consequence, if the density of free electrons is n , a current density \mathbf{J} will appear, which is given by Ohm's law

$$\mathbf{J} = -nev = \sigma \mathbf{E},$$

where

$$\sigma = \frac{ne^2\tau}{m_e}$$

is the conductivity, whose unit is Siemens per meter ($[\sigma] = \text{S/m}$). The current I going through a surface S is the flux of \mathbf{J} through S , and represents the total amount of charge per unit time crossing S

$$I = \Phi_{dS, \mathbf{J}} = \iint_S \mathbf{J}(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}')$$

- Since the electric charge is a conserved quantity, the charge density ρ and the current density \mathbf{J} are related at every point by the charge conservation law

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

- Ohm's law, in its macroscopic form states that the potential difference between points A and B of a conductor reads

$$V_A - V_B = R_{AB} I,$$

where I is the current flowing from A to B , and R_{AB} is the total resistance of the segment AB . For a rectilinear wire of section S and length l , the resistance is given by

$$R = \frac{\varrho l}{S},$$

where the resistivity ϱ is given by the inverse of the conductivity, $\varrho = 1/\sigma$.

- The relation between the resistance and the capacity of a conductor reads

$$RC = \frac{\epsilon_0}{\sigma},$$

where σ is the conductivity of the metal plates and ϵ_0 denotes the permittivity of vacuum. For a capacitor with a dielectric inserted between the two conductors, the relation remains valid if ϵ_0 is replaced by the permittivity of the dielectric.

- Hall effect: In the presence of a magnetic field \mathbf{B} , the electric field in a conductor carrying a steady state current density \mathbf{J} reads

$$\mathbf{E} = \rho\mathbf{J} + R_H \mathbf{B} \times \mathbf{J}$$

where the Hall coefficient is

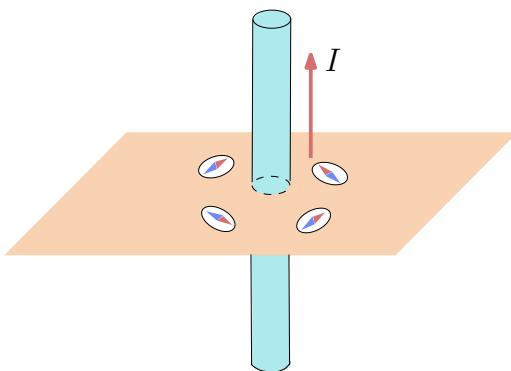
$$R_H = -1/ne.$$

Chapter 7

Magnetostatics

Introduction

Magnetic phenomena date back to ancient history. The first observations, in China and Greece, showed that an iron oxide called magnetite (Fe_3O_4) is what we call natural magnet, since it is capable of attracting other substances, such as small pieces of iron, which in turn acquire themselves magnetic properties after being in close proximity with magnetite. One of the first applications of magnetism occurred in navigation, thanks to compasses, which are made of a magnetized iron needle that orients itself in the direction of Earth's magnetic field. Since, many different magnetic materials formed by various mineral compounds have been identified, and their applications are countless. Magnets are widely employed in electronics, for example to store information in hard drives or to convert electricity into sound in speakers. They are used as well in magnetic nuclear resonance machines and are found inside electrical motors and electric generation plants.



In 1820, the Danish physicist Hans Christian Oersted discovered that electric currents are sources of magnetic fields. For example, the needle of a compass deviates in the presence of a current, confirming the close relationship between electricity and magnetism. Subsequently, the French physicists Biot, Savart and Ampère established the laws of magnetostatics, that is, of time-independent magnetic phenomena.

7.1 The magnetic field

Magnetic effects can be characterized through a magnetic field \mathbf{B} . As in electrostatics, the idea is to express the magnetic force acting on a charge q due to the influence of other charges as

$$\mathbf{F}_q = q\mathbf{v} \times \mathbf{B}(\mathbf{x}). \quad (7.1)$$

where \mathbf{v} denotes the velocity of charge q .

Definition 7.1: Magnetic field of a point charge

The magnetic field at \mathbf{x} generated by a charge q' of velocity \mathbf{v}' located at \mathbf{x}' is given by

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} q' \frac{\mathbf{v}' \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} \quad (7.2)$$

Note that if $\mathbf{E}(\mathbf{x})$ is the electric field generated by q' at the point \mathbf{x} , then we can write

$$\mathbf{B}(\mathbf{x}) = \mu_0 \epsilon_0 \mathbf{v}' \times \mathbf{E}(\mathbf{x}) = \frac{\mathbf{v}' \times \mathbf{E}}{c^2}. \quad (7.3)$$

It is clear from the definition that the magnetic field at a point \mathbf{x} generated by a moving charge is not a static field, since the vector $\mathbf{x} - \mathbf{x}'(t)$ depends on time. However, we will see later that a stationary electric current can generate static magnetic fields.

The S.I. unit for \mathbf{B} is the **Tesla** (symbol T), with $1\text{T} = 1\text{N A}^{-1}\text{m}$. The Gauss is also used, with $1\text{G} = 1 \times 10^{-4}\text{T}$, but does not belong to the SI unit system. The Earth's magnetic field has a magnitude of 0.2G , or $20\mu\text{T}$. As a comparison, an electromagnet can typically generate fields between 0.1 and 1T . A superconductor can generate fields of up to 50T .

7.1.1 The magnetic force

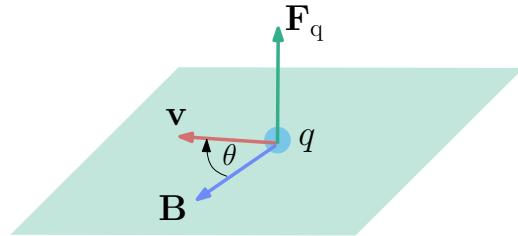
We see then from Eq. (7.1) that the magnetic force on a particle q of velocity \mathbf{v} in the presence of a magnetic field \mathbf{B} is

$$\mathbf{F}_q = q\mathbf{v} \times \mathbf{B}. \quad (7.4)$$

Comments

- The magnitude of the magnetic force is proportional to the magnitude $|\mathbf{v}|$ of the velocity, and to the sine of the angle θ between \mathbf{v} and \mathbf{B} . Indeed, $|\mathbf{F}| = qvB \sin \theta$.

- The magnetic force is perpendicular to the velocity \mathbf{v} . Consequently, it is a force that does not perform any work. A particle in a magnetic field only experiences a change in direction of its velocity, but the magnitude $|\mathbf{v}|$ of the latter will remain constant.



7.1.2 Where does the magnetic field of a magnet come from?

We see that there is a relationship between moving charges (currents) and magnetic fields. Why then does a material such as magnetite or an iron block, where no macroscopic electric current flows at equilibrium, generate magnetic fields? In a classical view of matter, every atom is composed of a positive nucleus and electrons revolving around it, which constitutes therefore a microscopic current. If we imagine that all these currents have the same orientation, they could generate a macroscopic magnetic field. However, this classic model is not satisfactory at atomic scales. The magnetism of a magnet does not come in general from an orbital motion of the electron, but from an intrinsic magnetism due to its spin, a property of every electron with no equivalent in classical mechanics. The magnetism of matter is then a quantum effect, and it will be discussed with more details in a devoted chapter.

7.1.3 Lorentz Force

Assuming that a particle of charge q is located in a region of space in which there is a magnetic field \mathbf{B} and an electric field \mathbf{E} , a net force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (7.5)$$

acts on the particle. This is called the **Lorentz Force**.

7.2 Field from a current - The Biot-Savart law (1820)

The French physicists Jean-Baptiste Biot and Félix Savart demonstrated experimentally in 1820 that the needle of a compass changes direction in the vicinity of a wire through which a current

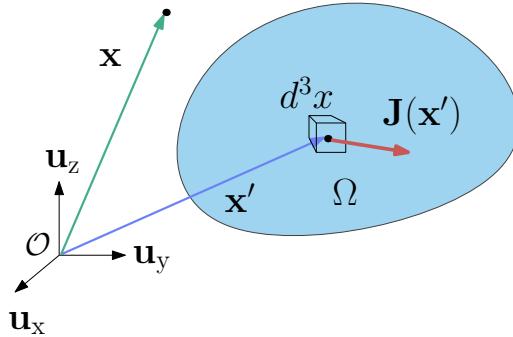
passes. By varying the distance between the wire and the compass needle, they managed to establish the law for the magnetic field generated by an arbitrary current. The Biot-Savart law is considered to be the fundamental law of magnetostatics, playing an equivalent role to that of Coulomb's law in electrostatics.

This discovery motivated many physicists to establish mathematical relationships between currents and magnetic fields. André-Marie Ampère was the first to establish a formula for the interaction force between two wires carrying currents. It was also Ampère who established the law, now known as Ampère's law, which plays in magnetostatics the role that Gauss's law has in electrostatics, and that will be presented below.

In the following, the Biot-Savart law will be deduced from the definition of the magnetic field generated by a moving point particle (see Eq. (7.2)).

7.2.1 Magnetic field from a continuous current distribution

Consider a fixed region of space Ω that encloses a distribution of charge density ρ . Charges are in motion within Ω , leading to a current density. We assume that the current density at any point of Ω is independent of time (stationary current). A volume element d^3x' around $\mathbf{x}' \in \Omega$ is then characterized by a total infinitesimal charge $dq(\mathbf{x}') = \rho(\mathbf{x}')d^3x'$ and a current density $\mathbf{J}(\mathbf{x}') = \rho(\mathbf{x}')\mathbf{v}(\mathbf{x}')$, where \mathbf{v} represents the velocity field.



From Eq. (7.2) and the superposition principle, the magnetic field at \mathbf{x} generated by the motion of charges in the volume d^3x' is given by

$$d\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} dq(\mathbf{x}') \mathbf{v}(\mathbf{x}') \times \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} = \frac{\mu_0}{4\pi} d^3x' \mathbf{J}(\mathbf{x}') \times \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \quad (7.6)$$

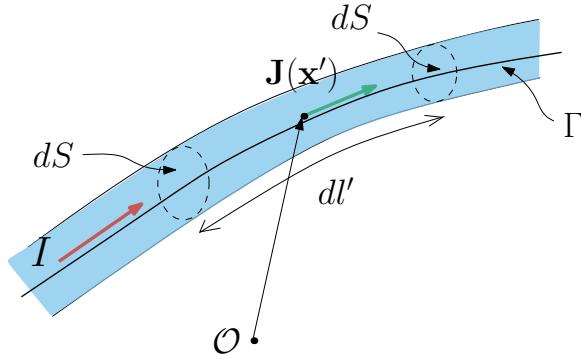
Finally, the total field at \mathbf{x} will be the integral over Ω given by

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'. \quad (7.7)$$

7.2.2 Magnetic field generated by a line current

Now consider a circuit with wires of small and constant section dS through which a uniform current I circulates. This means that we can represent the current flow as being concentrated on a one-dimensional path Γ , called a line current, that follows the shape of the circuit. Since the section dS is constant along the path, the modulus of the current density $|\mathbf{J}| = I/dS = J$ is constant as well. Consider a point \mathbf{x}' inside the conductor, and take the volume element $d^3x' = dl'(\mathbf{x}')dS$, where dl' is the differential line element at that point. We have

$$\mathbf{J}(\mathbf{x}')d^3x' = \mathbf{J}(\mathbf{x}')dSdl'(\mathbf{x}') = JdSdl(\mathbf{x}')\mathbf{t}(\mathbf{x}'),$$



where $\mathbf{t}(\mathbf{x}')$ denotes the direction of the current density, tangent to the curve Γ . Since $JdS = I$ is the current through the wire, by writing the vector line element $d\mathbf{l} = dl\mathbf{t}$ and using Eq. (7.6), we obtain the field generated by this current element $Id\mathbf{l}$:

$$d\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{Id\mathbf{l}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}.$$

Integrating over Γ , we get the expression of the Biot-Savart law for a line current.

Definition 7.2: The Biot-Savart law (1820)

A line current is described by a curve Γ through which a current I circulates. The magnetic field at \mathbf{x} generated by this circuit is given by the Biot-Savart law:

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 I}{4\pi} \int_{\Gamma} d\mathbf{l}(\mathbf{x}') \times \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}. \quad (7.8)$$

7.2.3 Magnetic field lines

A magnetic field lines is a continuous curve such that the tangent to the curve at every point \mathbf{x} coincides with the direction of $\mathbf{B}(\mathbf{x})$ at that point. Field lines can be found by writing

$$\mathbf{B}(\mathbf{x}) \times d\mathbf{x} = 0,$$

where $d\mathbf{x}$ is the infinitesimal displacement along the field line.

Reminder: If we know the expression for the components of the magnetic field as a function of spatial coordinates, the network of field lines can be obtained by solving an ordinary differential equation obtained from the condition $\mathbf{B}(\mathbf{x}) \times d\mathbf{x} = \mathbf{0}$. For instance in Cartesian coordinates, this condition is equivalent to

$$\frac{dx}{B_x(x, y, z)} = \frac{dy}{B_y(x, y, z)} = \frac{dz}{B_z(x, y, z)}.$$

Integrating by separation of variables, if possible, or by numerical methods otherwise, yields an equation for the field lines. Integration constants are fixed by initial conditions, i.e., by requesting that a field line passes through a given point.

The procedure is similar in other coordinate systems: For example, in cylindrical coordinates, the condition $\mathbf{B}(\mathbf{x}) \times d\mathbf{x} = \mathbf{0}$ reads

$$\frac{dr}{B_r(r, \theta, z)} = \frac{r d\theta}{B_\theta(r, \theta, z)} = \frac{dz}{B_z(r, \theta, z)}.$$

In magnetostatics, the magnetic field lines surround the electric currents that generate them and eventually close back on themselves, forming closed loops.

7.3 Magnetic fields for common current distributions

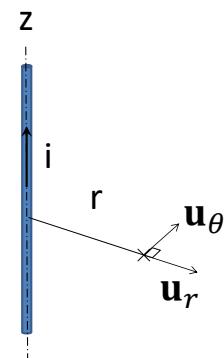
Suggested Exercises: Examples 7.10-7.12 in the exercise book.

It is very useful to remember the expressions for the magnetic field generated by common current distributions, as it facilitates the calculation of magnetic fields generated by other current distribution that can be simply seen as the superposition of common distributions. For example, a solenoid can be seen as a superposition of circular loops on the surface of a cylinder. The magnetic fields for the common distributions of currents are:

- Magnetic field at any point \mathbf{r} in space for an infinitely long wire carrying a current i .

In a cylindrical basis of coordinates with the wire along the z -axis:

$$\mathbf{B}(\mathbf{r}) = \mu_0 \frac{i}{2\pi r} \mathbf{u}_\theta.$$

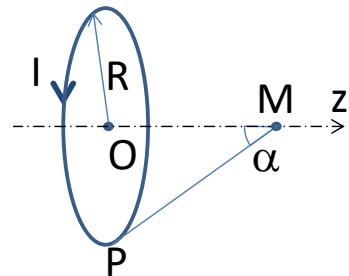


- Magnetic field at point M on the axis of a circular loop of radius R , carrying a current i .

In a cylindrical basis of coordinates with the axis of the loop along the z -axis:

$$\mathbf{B}(M) = \frac{\mu_0 N i}{2R} \sin^3 \alpha \mathbf{u}_z,$$

where α denotes the angle between the axis of the loop and line PM , the point P being any point of the loop.



- Magnetic field at point M on the axis of a solenoid of radius R , carrying a current i and having n loops per unit length.

In a cylindrical basis of coordinates with the axis of the solenoid along the z -axis:

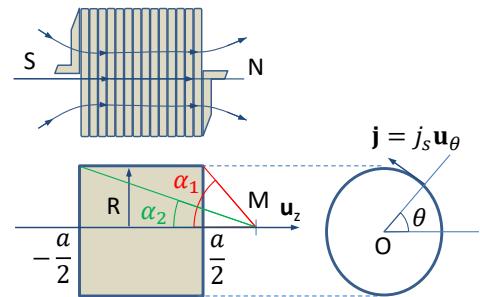
$$\mathbf{B}(M) = \frac{\mu_0 n i}{2} (\cos \alpha_2 - \cos \alpha_1) \mathbf{u}_z,$$

where α_1 and α_2 denote the angles between the axis of the solenoid and the lines joining point M and the edges of the solenoid.

The same formula is valid for a cylindrical sheet of surface current j_s , by replacing ni by j_s .

For an infinitely long solenoid, at any point M on the z -axis

$$\mathbf{B}(M) = \mu_0 n i \mathbf{u}_z.$$



7.4 Ampère's law

Ampère's law is one of the fundamental laws of magnetostatics. It plays a role equivalent to that of Gauss's law for electrostatics, i.e., it allows us to easily compute the magnetic field whenever the sources (here, the current density) exhibit a certain type of symmetry. In its differential form, Ampère's law states that a current density \mathbf{J} is the source of a rotational field \mathbf{B} .

7.4.1 Ampère's law - Differential form

From Biot-Savart's law (7.8), let us calculate the curl of an arbitrary magnetic field generated by a current density inside a volume $\Omega \subseteq \mathbb{R}^3$:

$$\nabla \times \mathbf{B}(\mathbf{x}) = \nabla \times \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' = \frac{\mu_0}{4\pi} \iiint_{\Omega} \nabla \times \left(\mathbf{J}(\mathbf{x}') \times \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \right) d^3x'.$$

Now we use the vector identity involving the two vectors $\mathbf{A}(\mathbf{x})$ and $\mathbf{C}(\mathbf{x})$:

$$\nabla \times (\mathbf{A} \times \mathbf{C}) = \mathbf{A}(\nabla \cdot \mathbf{C}) - \mathbf{C}(\nabla \cdot \mathbf{A}) + (\mathbf{C} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{C}.$$

By taking $\mathbf{A} = \mathbf{J}(\mathbf{x}')$, one has $\mathbf{C}(\nabla \cdot \mathbf{A}) = \mathbf{0}$ and $(\mathbf{C} \cdot \nabla)\mathbf{A} = \mathbf{0}$ since $\mathbf{J}(\mathbf{x}')$ is independent of \mathbf{x} . Choosing $\mathbf{C} = (\mathbf{x} - \mathbf{x}')/|\mathbf{x} - \mathbf{x}'|^3$ gives:

$$\nabla \times \mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \mathbf{J}(\mathbf{x}') \nabla \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' - \frac{\mu_0}{4\pi} \iiint_{\Omega} (\mathbf{J}(\mathbf{x}') \cdot \nabla) \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'.$$

In this last equality, the first term on the right hand side can be simplified by recognizing the divergence of the electric field generated by a point charge $q = 4\pi\epsilon_0$ at \mathbf{x}' . The differential form of Gauss's law then gives $\nabla \cdot (\mathbf{x} - \mathbf{x}')/|\mathbf{x} - \mathbf{x}'|^3 = 4\pi\delta(\mathbf{x} - \mathbf{x}')$. For the second term, of the form $(\mathbf{A} \cdot \nabla)\mathbf{C}$, we use the fact that \mathbf{C} is irrotational (meaning that it has zero curl) and that \mathbf{A} is independent of \mathbf{x} , then:

$$\nabla \times \mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} 4\pi \mathbf{J}(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}') d^3x' - \frac{\mu_0}{4\pi} \iiint_{\Omega} \nabla \left(\mathbf{J}(\mathbf{x}') \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \right) d^3x'.$$

If the current is stationary, $\nabla' \cdot \mathbf{J}(\mathbf{x}') = 0$ and therefore

$$\nabla \left(\mathbf{J}(\mathbf{x}') \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \right) = -\nabla' \left(\mathbf{J}(\mathbf{x}') \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \right).$$

Finally, using Green-Ostrogradsky's theorem, we obtain

$$\nabla \times \mathbf{B}(\mathbf{x}) = \mu_0 \iiint_{\Omega} \mathbf{J}(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}') d^3x' + \frac{\mu_0}{4\pi} \iint_{\partial\Omega} \left(\mathbf{J}(\mathbf{x}') \cdot \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} \right) \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}').$$

If $\mathbf{x} \in \Omega$, the first term on the right hand side equals $\mu_0 \mathbf{J}(\mathbf{x})$. If in addition the current density is zero in $\partial\Omega$ (the current density is localized inside the volume Ω), one obtains the differential form of Ampère's law.

Definition 7.3: Ampère's theorem (1826) - differential form

Let \mathbf{J} be a stationary ($\nabla \cdot \mathbf{J} = 0$) and localized current density inside a volume $\Omega \subseteq \mathbb{R}^3$, then:

$$\nabla \times \mathbf{B}(\mathbf{x}) = \mu_0 \mathbf{J}(\mathbf{x}).$$

The current density \mathbf{J} generates, at every point $\mathbf{x} \in \mathbb{R}^3$, a rotational magnetic field \mathbf{B} around \mathbf{J} .

7.4.2 Ampère's law - Integral form

Let Γ be a closed curve in \mathbb{R}^3 and $S(\Gamma)$ a surface bounded by the curve Γ . The Stokes theorem states:

$$\oint_{\Gamma} d\mathbf{x} \cdot \mathbf{B}(\mathbf{x}) = \iint_{S(\Gamma)} (\nabla \times \mathbf{B}(\mathbf{x})) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}).$$

The right hand side can be rewritten by using the differential form of Ampère's law:

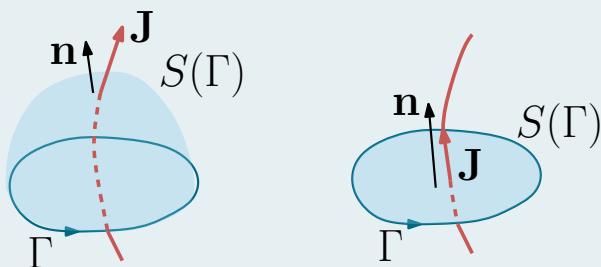
$$\oint_{\Gamma} d\mathbf{x} \cdot \mathbf{B}(\mathbf{x}) = \mu_0 \iint_{S(\Gamma)} \mathbf{J}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}) = \mu_0 \Phi_{S(\Gamma), \mathbf{J}},$$

showing that the right hand side is proportional to the flux of \mathbf{J} through $S(\Gamma)$, in other words, to the current passing through $S(\Gamma)$. To simplify the notation, this current will be written from now on as I_{Γ} .

Definition 7.4: Ampère's law (1826) - Integral form

The circulation of the magnetic field over a closed curve $\Gamma \in \mathbb{R}^3$ is proportional to the current $I(\Gamma)$ through any surface $S(\Gamma)$ whose bounding curve is Γ :

$$\oint_{\Gamma} d\mathbf{x} \cdot \mathbf{B}(\mathbf{x}) = \mu_0 I_{\Gamma}. \quad (7.9)$$



Remarks

1. Here, a proper orientation of both Γ and $S(\Gamma)$ is assumed as required by Stokes's theorem. If the surface $S(\Gamma)$ is looked in such a way that the normal vector is pointing towards you, then the curve Γ should be oriented counterclockwise.
2. Given a curve Γ , an infinity of surfaces $S(\Gamma)$ can be chosen that are bounded by Γ . Remarkably, the flux of \mathbf{J} through $S(\Gamma)$ does not depend on this choice. This is due charge conservation in magnetostatics, i.e., the fact that $\nabla \cdot \mathbf{J} = 0$, as will be shown later. In the figure below, two different possibilities are shown, the current passing through $S(\Gamma)$ being the same in both cases.

7.5 Symmetry arguments in magnetostatics

Ampère's law is particularly useful whenever the sources of the magnetic field have some spatial symmetries. According to Curie's principle, in order to determine the symmetries of the magnetic field (effect) it is sufficient to study the symmetries of the current density \mathbf{J} (cause).

- **INVARIANCE BY SPATIAL TRANSLATION ALONG AN AXIS**

Consider the case of an invariant current density under translation along an axis (Oz for example). This means, in Cartesian coordinates:

$$\mathbf{J}(x, y, z + a) = \mathbf{J}(x, y, z) \quad \forall a,$$

so that the current density is independent of the z -coordinate, $\mathbf{J} = \mathbf{J}(x, y)$, and this translates directly into an invariance of the magnetic field field upon this coordinate:

$$\mathbf{B}(x, y, z) = \mathbf{B}(x, y).$$

- **INVARIANCE BY SPATIAL ROTATION AROUND AN AXIS**

Consider the case of a current density that is invariant under rotation around the z -axis. This means, in cylindrical coordinates:

$$\mathbf{J}(r, \theta + a, z) = \mathbf{J}(r, \theta, z) \quad \forall a,$$

so that the current density is independent of θ , i.e., $\mathbf{J}(r, \theta, z) = \mathbf{J}(r, z)$, and consequently, the magnetic field does not depend on this coordinate either:

$$\mathbf{B}(r, \theta, z) = \mathbf{B}(r, z).$$

- MIRROR SYMMETRY

To constrain the orientation of the magnetic field, one needs to determine whether any mirror symmetry is present. In order to demonstrate the mirror symmetries in magnetostatics, one proceeds in a similar way as in electrostatics. Any plane Π allows us to partition \mathbb{R}^3 into two disjoint regions D_1 and D_2 . As such, any point $\mathbf{u}' \in D_2$ can be written as the reflection of a point $\mathbf{u} \in D_1$ in the (mirror) plane Π . Then, the magnetic field at position \mathbf{x} can be written as an integral over D_1 only, according to Biot-Savart's law:

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{D_1} \mathbf{J}(\mathbf{u}) \times \frac{\mathbf{x} - \mathbf{u}}{|\mathbf{x} - \mathbf{u}|^3} d^3 u + \frac{\mu_0}{4\pi} \iiint_{D_2} \mathbf{J}(\mathbf{u}') \times \frac{\mathbf{x} - \mathbf{u}'}{|\mathbf{x} - \mathbf{u}'|^3} d^3 u'.$$

Similarly, the magnetic field at $\mathbf{x}' = \text{sym}_\Pi \mathbf{x}$ is written as:

$$\mathbf{B}(\text{sym}_\Pi \mathbf{x}) = \mathbf{B}(\mathbf{x}') = \frac{\mu_0}{4\pi} \iiint_{D_1} \mathbf{J}(\mathbf{u}) \times \frac{\mathbf{x}' - \mathbf{u}}{|\mathbf{x}' - \mathbf{u}|^3} d^3 u + \frac{\mu_0}{4\pi} \iiint_{D_2} \mathbf{J}(\mathbf{u}') \times \frac{\mathbf{x}' - \mathbf{u}'}{|\mathbf{x}' - \mathbf{u}'|^3} d^3 u'.$$

Noting that $|\mathbf{x} - \mathbf{u}| = |\mathbf{x}' - \mathbf{u}'|$, $|\mathbf{x} - \mathbf{u}'| = |\mathbf{x}' - \mathbf{u}|$ and that $\mathbf{a} - \mathbf{b} = \mathbf{a}' - \mathbf{b}'$ for any pair of vectors \mathbf{a} and \mathbf{b} :

$$\mathbf{B}(\text{sym}_\Pi \mathbf{x}) = \mathbf{B}(\mathbf{x}') = \frac{\mu_0}{4\pi} \iiint_{D_1} \mathbf{J}(\mathbf{u}) \times \frac{(\mathbf{x} - \mathbf{u}')'}{|(\mathbf{x} - \mathbf{u}')'|^3} d^3 u + \frac{\mu_0}{4\pi} \iiint_{D_2} \mathbf{J}(\mathbf{u}') \times \frac{(\mathbf{x} - \mathbf{u}')'}{|(\mathbf{x} - \mathbf{u}')'|^3} d^3 u'.$$

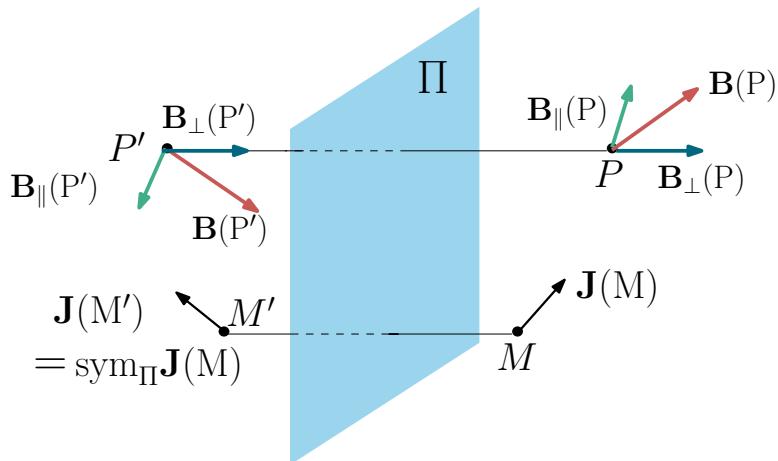
From here, it is clear that:

1. If \mathbf{J} has mirror symmetry with respect to Π , then for any vector \mathbf{b} one has $\mathbf{J}(\mathbf{u}') \times \mathbf{b}' = -\text{sym}_\Pi \mathbf{J}(\mathbf{u}) \times \mathbf{b}$, and it follows that \mathbf{B} has mirror antisymmetry with respect to Π :

$$\mathbf{B}(\mathbf{x}') = \frac{\mu_0}{4\pi} \iiint_{D_1} -\left(\mathbf{J}(\mathbf{u}') \times \frac{(\mathbf{x} - \mathbf{u}')'}{|(\mathbf{x} - \mathbf{u}')'|^3} \right)' d^3 u + \frac{\mu_0}{4\pi} \iiint_{D_2} -\left(\mathbf{J}(\mathbf{u}) \times \frac{(\mathbf{x} - \mathbf{u})}{|\mathbf{x} - \mathbf{u}|^3} \right)' d^3 u'.$$

Then

$\text{Mirror symmetry for } \mathbf{J} \rightarrow \mathbf{B}(\mathbf{x}') = -\text{sym}_\Pi \mathbf{B}(\mathbf{x}).$



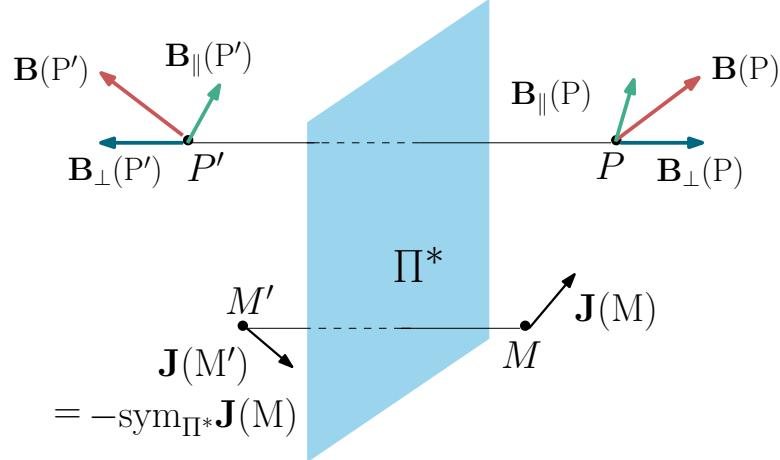
The mirror symmetry of \mathbf{J} implies $\mathbf{B}_\perp(P) = \mathbf{B}_\perp(P')$ and $\mathbf{B}_\parallel(P) = -\mathbf{B}_\parallel(P')$. In particular, for a point $P \in \Pi$, one has $P' = P$ and therefore $\mathbf{B}_\parallel(P) = -\mathbf{B}_\parallel(P) = 0$. In other words, for a point P inside a symmetry plane Π , the magnetic field is perpendicular to the plane of symmetry.

2. If \mathbf{J} has mirror antisymmetry with respect to Π^* , then for any vector \mathbf{b} , one has $\mathbf{J}(\mathbf{u}') \times \mathbf{b}' = \text{sym}_{\Pi^*}\mathbf{J}(\mathbf{u}) \times \mathbf{b}$, and it follows that \mathbf{B} has mirror symmetry with respect to Π^* :

$$\mathbf{B}(\mathbf{x}') = \frac{\mu_0}{4\pi} \iiint_{D_1} \left(\mathbf{J}(\mathbf{u}') \times \frac{(\mathbf{x} - \mathbf{u}')}{|\mathbf{x} - \mathbf{u}'|^3} \right)' d^3u + \frac{\mu_0}{4\pi} \iiint_{D_2} \left(\mathbf{J}(\mathbf{u}) \times \frac{(\mathbf{x} - \mathbf{u})}{|\mathbf{x} - \mathbf{u}|^3} \right)' d^3u'.$$

Then

$\text{Mirror antisymmetry for } \mathbf{J} \rightarrow \mathbf{B}(\mathbf{x}') = \text{sym}_{\Pi^*}\mathbf{B}(\mathbf{x}).$



This mirror antisymmetry of \mathbf{J} implies $\mathbf{B}_\perp(P) = -\mathbf{B}_\perp(P')$ and $\mathbf{B}_\parallel(P) = \mathbf{B}_\parallel(P')$. In particular, for a point $P \in \Pi^*$, one has $P' = P$ and therefore $\mathbf{B}_\perp(P) = -\mathbf{B}_\perp(P) = \mathbf{0}$. In other words, for any point P inside an antisymmetry plane Π^* , the magnetic field is within the plane of antisymmetry.

Remark: The magnetic field is a pseudo-vector, it gets a minus sign upon reflection due to its expression involving a cross product.

Suggested exercises on Ampère's law: Ex. 8.1-8.6 in the exercise book

7.6 The vector potential

We recall from electrostatics that the electric potential V can be defined by the relation $\mathbf{E} = -\nabla V$. Similarly in magnetostatics, we will see that it is possible to define a vector potential \mathbf{A} such that the magnetic field can be obtained as the curl of this potential,

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

7.6.1 Vector potential of a point charge

Let us suppose that at a given instant, a particle of charge q and velocity \mathbf{v} is at position \mathbf{x}' . From Eq. (7.2), the magnetic field at \mathbf{x} due to this charge is

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} q \mathbf{v} \times \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}.$$

Considering now that $\frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = -\nabla \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right)$, the magnetic field writes

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \left(\nabla \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \times q \mathbf{v} \right).$$

Since $q \mathbf{v}$ is independent of \mathbf{x} , using the vectorial calculus identity

$$\nabla \times (a \mathbf{A}) = a (\nabla \times \mathbf{A}) + \nabla a \times \mathbf{A},$$

where a is a scalar and \mathbf{A} a vector, we can write $\nabla \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \times q \mathbf{v} = \nabla \times \left(\frac{q \mathbf{v}}{|\mathbf{x} - \mathbf{x}'|} \right)$ and finally:

$$\mathbf{B}(\mathbf{x}) = \nabla \times \left(\frac{\mu_0}{4\pi} \frac{q \mathbf{v}}{|\mathbf{x} - \mathbf{x}'|} \right).$$

Definition 7.5: Vector potential of a point charge

The vector potential generated by a point charge q at \mathbf{x}' , of velocity \mathbf{v} , is given by:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{q \mathbf{v}}{|\mathbf{x} - \mathbf{x}'|}$$

(7.10)

and the magnetic field generated by this charge verifies:

$$\mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x}).$$

7.6.2 Vector potential of an arbitrary current distribution

The vector potential can be generalized for an arbitrary current distribution. Indeed, Biot-Savart's law can be rewritten:

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \mathbf{J}(\mathbf{x}') \times \nabla \left(-\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) d^3 x' = \nabla \times \left(\frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3 x' \right).$$

Definition 7.6: Vector potential of an arbitrary current distribution

For an arbitrary current density localized inside a volume $\Omega \subseteq \mathbb{R}^3$, the vector potential is given by the integral:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (7.11)$$

For the case of a line of current, this reduces to:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0 I}{4\pi} \int_{\Gamma} \frac{d\mathbf{l}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (7.12)$$

7.6.3 The vector potential is not unique

Even if Eq. (7.11) defines a unique vector potential, one can define in reality an infinity of vector potentials satisfying

$$\nabla \times \mathbf{A} = \mathbf{B}. \quad (7.13)$$

Indeed, if \mathbf{A} is defined from Eq. (7.11), one can always add to \mathbf{A} the gradient of an arbitrary scalar field:

$$\mathbf{A}' = \mathbf{A} + \nabla\psi.$$

Since $\nabla \times \nabla\psi = \mathbf{0}$, both \mathbf{A} and \mathbf{A}' define the same magnetic field:

$$\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times \mathbf{A}'.$$

We know from the Helmholtz theorem that to uniquely determine a vector field, one must specify its curl and its divergence. The vector potential defined by Eq. (7.11) satisfies

$$\nabla \times \mathbf{A} = \mathbf{B} \quad \text{and} \quad \nabla \cdot \mathbf{A} = 0,$$

which corresponds to a particular choice among all the possible vector fields that satisfy $\nabla \times \mathbf{A} = \mathbf{B}$. This choice corresponds to what is called the Coulomb gauge.

7.6.4 Poisson's equation for the vector potential

In a similar way that the electrostatic potential verifies Poisson's equation, it is possible to demonstrate that the vector potential in the Coulomb gauge is a solution to

$$\nabla^2 \mathbf{A}(\mathbf{x}) = -\mu_0 \mathbf{J}(\mathbf{x}). \quad (7.14)$$

In other words, the i^{th} -component of \mathbf{A} satisfies Poisson's equation with a source term given by $-\mu_0 J_i$. Indeed, by using the vector identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, and the definition of the vector potential $\nabla \times \mathbf{A} = \mathbf{B}$, we obtain

$$\nabla \times \mathbf{B} = \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}.$$

Using Ampère's law and the fact that the vector potential in Coulomb's gauge has zero divergence $\nabla \cdot \mathbf{A} = 0$, then we obtain Eq. (7.14) by identifying the left hand side using Maxwell-Ampère's equation

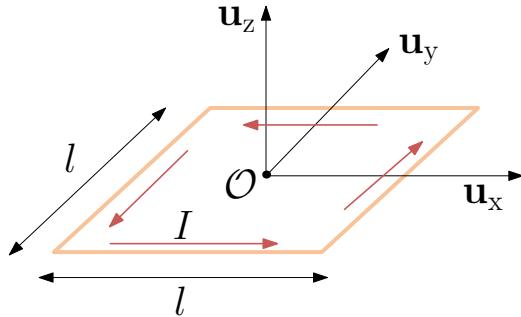
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}.$$

7.7 The magnetic dipole

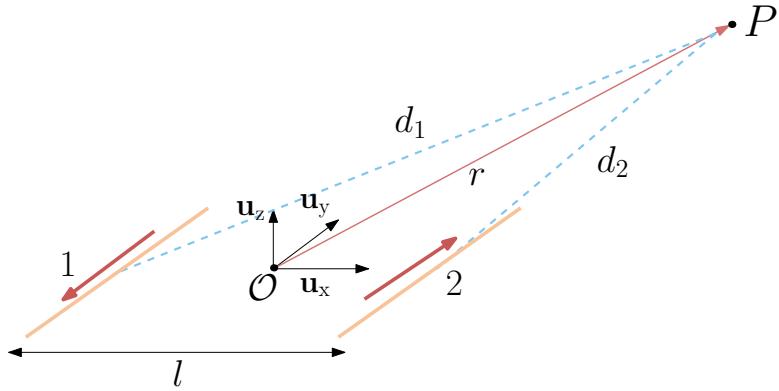
It is possible to show that the magnetic field in the axis of a circular current loop decays as $1/r^3$ sufficiently far from the loop (see Ex. 7.11 in the exercise book). This is reminiscent of the $1/r^3$ dependence of the electric field generated by an electric dipole. In magnetostatics, a small loop of current defines a magnetic dipole. We will show in this section that the field lines of a magnetic dipole are equivalent to those of an electric dipole.

7.7.1 Magnetic field of a magnetic dipole

Consider a square loop of size l carrying a current I in the Oxy plane in the counterclockwise direction. Let us find the magnetic field at distances much larger than the typical size of the loop.



First, we compute the vector potential at a distance r from the origin (point P) generated by the superposition of the four current elements composing the loop.



The two segments parallel to the y -axis give:

$$\mathbf{A}_1(\mathbf{x}) \approx \frac{\mu_0 I}{4\pi} \frac{(-l)}{d_1} \mathbf{u}_y, \quad \mathbf{A}_2(\mathbf{x}) \approx \frac{\mu_0 I}{4\pi} \frac{l}{d_2} \mathbf{u}_y,$$

where

$$d_1 = \sqrt{(x + l/2)^2 + y^2 + z^2} = \sqrt{r^2 + lx + l^2/4},$$

$$d_2 = \sqrt{(x - l/2)^2 + y^2 + z^2} = \sqrt{r^2 - lx + l^2/4}.$$

From here, the y -component of the vector potential is given by the superposition

$$A_y(\mathbf{x}) = \frac{\mu_0 I l}{4\pi} \left(\frac{1}{d_2} - \frac{1}{d_1} \right) \approx \frac{\mu_0 I l^2 x}{4\pi r^3}$$

where the first order in x/r^2 has been kept for the approximation. A similar calculation yields, for the x component:

$$A_x(\mathbf{x}) \approx -\frac{\mu_0 I l^2 y}{4\pi r^3}.$$

Recalling that $S = l^2$ is the planar surface enclosed by the loop, one obtains finally,

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0 I S}{4\pi r^3} (x \mathbf{u}_y - y \mathbf{u}_x).$$

Defining the magnetic moment of the loop as $\mathbf{M} = IS\mathbf{n}$ where n is the oriented normal to the plane surface enclosed by the loop, the vector potential can be rewritten as:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{r}}{r^3}.$$

The magnetic field far from the magnetic moment is then obtained as

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) = \nabla \times \left(\frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{r}}{r^3} \right).$$

Now we use the following identity $\nabla \times (\mathbf{m} \times \mathbf{r}/r^3) = \mathbf{m}(\nabla \cdot \mathbf{r}/r^3) - (\mathbf{m} \cdot \nabla)\mathbf{r}/r^3$, which can be simplified considering that the divergence of \mathbf{r}/r^3 is zero for $r \neq 0$ and that \mathbf{m} is a constant vector:

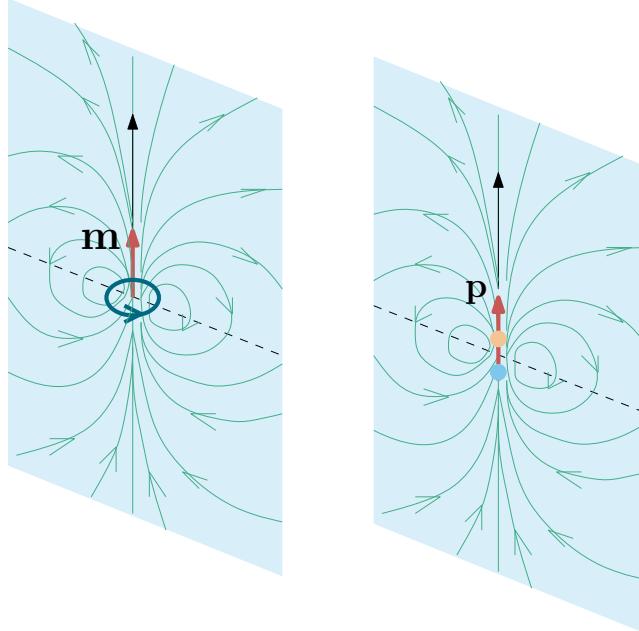
$$\nabla \times \left(\frac{\mathbf{m} \times \mathbf{r}}{r^3} \right) = -\nabla \left(\mathbf{m} \cdot \frac{\mathbf{r}}{r^3} \right)$$

and one obtains the remarkable result

$$\mathbf{B}(\mathbf{r}) = -\nabla \left\{ \frac{\mu_0}{4\pi} \frac{\mathbf{m} \cdot \mathbf{r}}{r^3} \right\} = -\nabla \psi,$$

where the scalar field ψ is identical to the electric potential of an electric dipole moment $\mathbf{p} = \mathbf{m}$, after changing μ_0 into $1/\epsilon_0$ (See chapter 2). One concludes that, for $\mathbf{x} \neq \mathbf{0}$, the magnetic field of a magnetic dipole has an identical expression as the electric field generated by an electric dipole, which we recall here

$$\mathbf{E}(\mathbf{r}) = -\nabla \left\{ \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p} \cdot \mathbf{r}}{r^3} \right\}$$



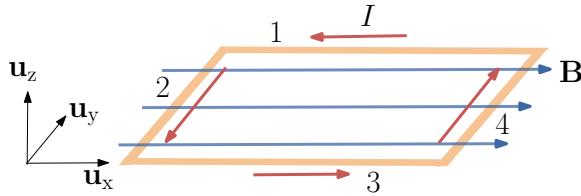
7.7.2 Magnetic moment

In electrostatics, we saw that a system of two charges of equal magnitude and opposite sign is completely described by its dipole moment \mathbf{p} . We have also seen that a dielectric medium can be viewed as a collection of microscopic dipole moments, and that the medium is completely characterized by its polarization \mathbf{P} , which represents the dipole moment per unit volume. In the presence of an electric field \mathbf{E} , a dipole feels a force $\mathbf{F} = \nabla(\mathbf{p} \cdot \mathbf{E})$ and a torque $\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}$. Similarly, we will define the magnetic moment \mathbf{m} of a closed current so that the force and torque on it in the presence of a magnetic field are written

$$\begin{aligned}\mathbf{F} &= \nabla(\mathbf{m} \cdot \mathbf{B}), \\ \boldsymbol{\tau} &= \mathbf{m} \times \mathbf{B}.\end{aligned}$$

• TORQUE ON A LOOP

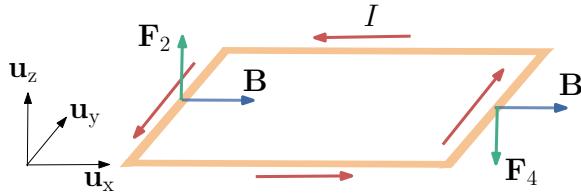
Before generalizing the definition of a dipole moment and the expressions for the magnetic force and torque exerted on it, let's see what happens when we place a rectangular loop that carries a stationary current I in the plane Oxy , in the presence of a uniform magnetic field $\mathbf{B}(\mathbf{x}) = B\mathbf{u}_x$ parallel to the plane of the loop.



We see that the magnetic forces acting on segments 1 and 3 are null because the vectors $\mathbf{l}_1 = -b\mathbf{u}_x$ and $\mathbf{l}_3 = b\mathbf{u}_x$ are parallel to the magnetic field \mathbf{B} . On the other hand, the forces acting on segments 2 and 4 are

$$\begin{aligned}\mathbf{F}_2 &= I(-a\mathbf{u}_y) \times B\mathbf{u}_x = Iab\mathbf{u}_z, \\ \mathbf{F}_4 &= I(a\mathbf{u}_y) \times B\mathbf{u}_x = -Iab\mathbf{u}_z.\end{aligned}$$

From here, it is clear that the net force is zero, as expected for a closed current in the presence of a uniform magnetic field. However, the forces \mathbf{F}_2 and \mathbf{F}_4 will produce a torque that generates a rotation of the loop with respect to the y -axis.



The torque with respect to the center of the loop is

$$\begin{aligned}\boldsymbol{\tau} &= \left(-\frac{b}{2}\mathbf{u}_x\right) \times \mathbf{F}_2 + \left(\frac{b}{2}\mathbf{u}_x\right) \times \mathbf{F}_4 = \left(-\frac{b}{2}\right)\mathbf{u}_x \times (Iab\mathbf{u}_z) + \left(\frac{b}{2}\mathbf{u}_x\right) \times (-Iab\mathbf{u}_z) \\ &= \left(\frac{IabB}{2} + \frac{IabB}{2}\right)\mathbf{u}_y = IabB\mathbf{u}_y = ISB\mathbf{u}_y,\end{aligned}$$

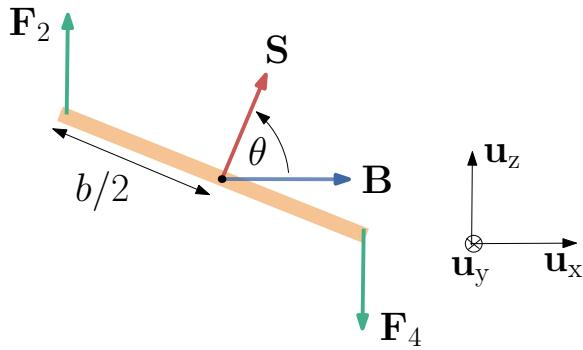
where $S = ab$ represents the surface of the loop. It is convenient to introduce the surface vector

$$\mathbf{S} = S\mathbf{n},$$

with \mathbf{n} the unit vector in the normal direction to the plane of the loop. In this case, $\mathbf{n} = \mathbf{u}_z$. The expression for torque can then be rewritten as

$$\boldsymbol{\tau} = IS\mathbf{n} \times \mathbf{B}.$$

Consider now the more general case where the loop forms an angle θ with respect to the magnetic field.



From the figure,

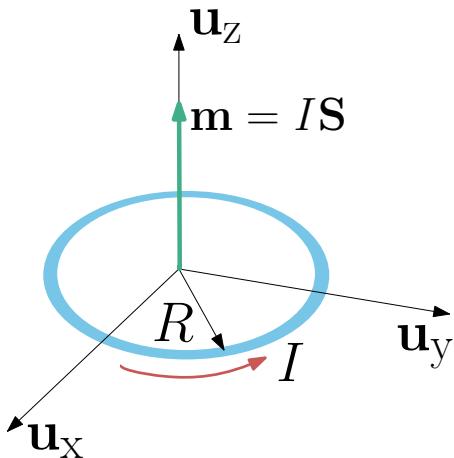
$$\mathbf{r}_2 = \frac{b}{2} (-\sin \theta \mathbf{u}_x + \cos \theta \mathbf{u}_z) = -\mathbf{r}_4$$

and the net torque is

$$\begin{aligned}\boldsymbol{\tau} &= \mathbf{r}_2 \times \mathbf{F}_2 + \mathbf{r}_4 \times \mathbf{F}_4 = 2\mathbf{r}_2 \times \mathbf{F}_2 = 2\frac{b}{2} (-\sin \theta \mathbf{u}_x + \cos \theta \mathbf{u}_z) \times (IaB \mathbf{u}_z) \\ &= IabB \sin \theta \mathbf{u}_y = I\mathbf{S} \times \mathbf{B}.\end{aligned}$$

This is the same expression as obtained above. Note that the torque is null when the normal to the loop is oriented in the direction parallel to the magnetic field. The vector $I\mathbf{S}$ is, by definition, the magnetic moment of the loop:

$$\mathbf{m} = I\mathbf{S}.$$



The direction of \mathbf{m} is the same as that of the surface vector \mathbf{S} (perpendicular to the plane of the loop). With this, the torque acting on a current loop is

$$\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}.$$

A loop in a magnetic field will rotate until its magnetic moment is completely aligned with the field. Note the analogy with the torque exerted by an electric field on a dipole.

Definition 7.7: Magnetic moment

The magnetic moment \mathbf{m} of an arbitrary current I in a circuit defined by the closed curve Γ (not necessarily plane) is given by:

$$\mathbf{m} = \frac{I}{2} \oint_{\Gamma} \mathbf{x} \times d\mathbf{x}. \quad (7.15)$$

Let $S(\Gamma)$ be any oriented surface whose bounding curve is Γ . The magnetic moment is then equivalently expressed as a function of the surface vector \mathbf{S} .

$$\mathbf{m} = IS = I \iint_{S(\Gamma)} \mathbf{n}(\mathbf{x}) dS. \quad (7.16)$$

The result does not depend on the choice of the surface provided it is bounded by Γ . For a plane current, Eq. (7.16) is reduced to $\mathbf{m} = IS\mathbf{n}$, with S the surface area enclosed by the current and \mathbf{n} the normal vector to the plane which contains the current (if the current is in the plane xy , $\mathbf{n} = \mathbf{u}_z$ if the current circulates counterclockwise).

More generally, the magnetic moment of an arbitrary distribution of current \mathbf{J} contained in a volume Ω is given by:

$$\mathbf{m} = \frac{1}{2} \iiint_{\Omega} \mathbf{x} \times \mathbf{J}(\mathbf{x}) d^3x.$$

In the international system, the unit of magnetic moment is A m^2 .

Proof: To prove that Eq. (7.15) and Eq. (7.16) are equivalent, we need to show that the surface vector

$$\mathbf{S} = \iint_{S(\Gamma)} \mathbf{n}(\mathbf{x}) dS = \frac{1}{2} \oint_{\Gamma} \mathbf{x} \times d\mathbf{x}. \quad (7.17)$$

It is sufficient to prove that both sides are equal component-wise. Let us consider the projection $\mathbf{S} \cdot \mathbf{u}$ of the surface vector \mathbf{S} on a unit vector \mathbf{u} of a Cartesian basis. Since \mathbf{u} is uniform over Γ , it can be moved inside the integral and we can then apply a circular permutation of the triple product, which writes

$$\mathbf{S} \cdot \mathbf{u} = \frac{1}{2} \oint_{\Gamma} (\mathbf{x} \times d\mathbf{x}) \cdot \mathbf{u} = \frac{1}{2} \oint_{\Gamma} (\mathbf{u} \times \mathbf{x}) \cdot d\mathbf{x}.$$

Now Ostrogradsky's theorem allows us to transform the contour integral into an integral over a surface bounding Γ

$$\mathbf{S} \cdot \mathbf{u} = \frac{1}{2} \iint_{S(\Gamma)} \nabla \times (\mathbf{u} \times \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS.$$

Any constant vector \mathbf{u} can be expressed as

$$\mathbf{u} = \frac{1}{2} \nabla \times (\mathbf{u} \times \mathbf{x}).$$

Thus,

$$\mathbf{S} \cdot \mathbf{u} = \iint_{S(\Gamma)} \mathbf{u} \cdot \mathbf{n}(\mathbf{x}) dS = \left(\iint_{S(\Gamma)} \mathbf{n}(\mathbf{x}) dS \right) \cdot \mathbf{u},$$

which is the dot product of \mathbf{u} with the vector surface on the left-hand-side of (7.17). Equation 7.17 is thus proved component-wise and the fact that its right-hand-side does not depend on the surface S but only on the contour Γ shows that the choice of the surface does not affect the dipole moment.

7.7.3 Multipolar expansion for the magnetic field

Let us assume a current distribution localized in a bounded volume Ω . The vector potential is given by:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'.$$

For a point \mathbf{x} far away from Ω , $|\mathbf{x}'| \ll |\mathbf{x}|$, a small- $|\mathbf{x}'|/|\mathbf{x}|$ Taylor expansion at first order reads

$$|\mathbf{x} - \mathbf{x}'|^{-1} = \left(|\mathbf{x}|^2 + |\mathbf{x}'|^2 - 2\mathbf{x} \cdot \mathbf{x}' \right)^{-1/2} = \frac{1}{|\mathbf{x}|} \left(1 + \frac{|\mathbf{x}'|^2}{|\mathbf{x}|^2} - \frac{2\mathbf{x} \cdot \mathbf{x}'}{|\mathbf{x}|^2} \right)^{-1/2} \approx \frac{1}{|\mathbf{x}|} + \frac{\mathbf{x} \cdot \mathbf{x}'}{|\mathbf{x}|^3},$$

so that the vector potential far away from the current distribution writes

$$\mathbf{A}(\mathbf{x}) \approx \frac{\mu_0}{4\pi|\mathbf{x}|} \iiint_{\Omega} \mathbf{J}(\mathbf{x}') d^3x' + \frac{\mu_0}{4\pi|\mathbf{x}|^3} \iiint_{\Omega} \mathbf{J}(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) d^3x'.$$

In this expansion, the first term (that would correspond to a magnetic monopole) is zero. To demonstrate this, it is sufficient to show that each component is zero, i.e., $\iiint_{\Omega} J_i(\mathbf{x}') d^3x' = 0$ for $i = x, y, z$. Consider the product rule differentiation formula $\nabla \cdot (f\mathbf{J}) = \mathbf{J} \cdot \nabla f + f \nabla \cdot \mathbf{J}$, where f is a scalar quantity. For a stationary current, $\nabla \cdot \mathbf{J} = 0$, leading to $\mathbf{J} \cdot \nabla f = \nabla \cdot (\mathbf{J}f)$. Integrating over Ω yields

$$\iiint_{\Omega} (\mathbf{J} \cdot \nabla f) d^3x = \iiint_{\Omega} \nabla \cdot (\mathbf{J}f) d^3x$$

and Green-Ostrogradsky's theorem can be used to transform the right-hand side:

$$\iiint_{\Omega} (\mathbf{J} \cdot \nabla f) d^3x = \iint_{\partial\Omega} \mathbf{J}f \cdot \mathbf{n} dS = 0,$$

which is zero since the current is localized inside Ω . Finally, taking $f(\mathbf{x}) = x$,

$$\iiint_{\Omega} \mathbf{J}(\mathbf{x}) \cdot \nabla x d^3x = \iiint_{\Omega} J_x(\mathbf{x}) d^3x = 0.$$

In the same way, the components along y or z are shown to be equally zero by using $f(\mathbf{x}) = y$ and $\mathbf{J}(\mathbf{x}) \cdot \nabla y = J_y(\mathbf{x})$, or $f(\mathbf{x}) = z$ and $\mathbf{J}(\mathbf{x}) \cdot \nabla z = J_z(\mathbf{x})$. Then,

$$\iiint_{\Omega} \mathbf{J}(\mathbf{x}) d^3x = \mathbf{0}.$$

Let us now work on each component of the dominant term of the vector potential, i.e., the term proportional to the vector of components $\iiint_{\Omega} J_i(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) d^3x'$ for $i = x, y, z$. Writing $J_i(\mathbf{x}') = \mathbf{J}(\mathbf{x}') \cdot \nabla' x'_i$, and using a the repeated index indicates a sum (Einstein's notation) in $\mathbf{x}' \cdot \mathbf{x} = \sum_k x'_k x_k = x'_k x_k$, we find

$$\begin{aligned} J_i(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) &= (\mathbf{J}(\mathbf{x}') \cdot \nabla' x'_i) (x'_k x_k) \\ &= \{\mathbf{J}(\mathbf{x}') \cdot (x'_k \nabla' x'_i)\} x_k \\ &= \{\mathbf{J}(\mathbf{x}') \cdot (\nabla' (x'_k x'_i) - x'_i \nabla' x'_k)\} x_k \\ &= \{\nabla' \cdot (x'_i x'_k \mathbf{J}(\mathbf{x}')) - x'_i \underbrace{\mathbf{J}(\mathbf{x}') \cdot (\nabla' x'_k)}_{J_k(\mathbf{x}')} \} x_k \\ &= \nabla' \cdot (x'_i x'_k \mathbf{J}(\mathbf{x}')) x_k - x'_i (\mathbf{J}(\mathbf{x}') \cdot \mathbf{x}), \end{aligned}$$

where we used the product rule differentiation formula and $\nabla' \cdot \mathbf{J}(\mathbf{x}') = 0$ to transform the first term on the right-hand side of the third line into a divergence. Integrating the resulting equation over Ω and eliminating the divergence term by the Green-Ostrogradsky theorem and the fact that $J(\mathbf{x}') = \mathbf{0}$ on $\partial\Omega$, we find, for $i = x, y, z$,

$$\iiint_{\Omega} J_i(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) d^3x' = \iiint_{\Omega} x'_i (\mathbf{J}(\mathbf{x}') \cdot \mathbf{x}) d^3x',$$

that is, gathering all components,

$$\iiint_{\Omega} \mathbf{J}(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) d^3x' = \iiint_{\Omega} \mathbf{x}' (\mathbf{J}(\mathbf{x}') \cdot \mathbf{x}) d^3x'.$$

Finally, using the triple product identity $(\mathbf{x}' \times \mathbf{J}(\mathbf{x}')) \times \mathbf{x} = (\mathbf{x}' \cdot \mathbf{x}) \mathbf{J}(\mathbf{x}') - (\mathbf{x} \cdot \mathbf{J}(\mathbf{x}')) \mathbf{x}'$ and integrating in over Ω , we find

$$\iiint_{\Omega} \mathbf{J}(\mathbf{x}') (\mathbf{x}' \cdot \mathbf{x}) d^3x' = \frac{1}{2} \iiint_{\Omega} (\mathbf{x}' \times \mathbf{J}(\mathbf{x}')) \times \mathbf{x} d^3x'$$

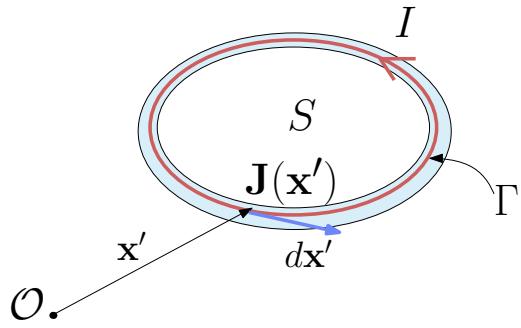
and the following result is obtained for the vector potential far from an arbitrary current distribution:

$$\mathbf{A}(\mathbf{x}) \approx \frac{\mu_0}{4\pi} \times \underbrace{\frac{1}{2} \left(\iiint_{\Omega} \mathbf{x}' \times \mathbf{J}(\mathbf{x}') d^3x' \right)}_{\mathbf{m}} \times \frac{\mathbf{x}}{|\mathbf{x}|^3}.$$

One recognizes the magnetic moment of an arbitrary current distribution:

$$\mathbf{m} = \iiint_{\Omega} \mathbf{x}' \times \mathbf{J}(\mathbf{x}') d^3x'.$$

Indeed, for the case of a current I carried by a planar loop Γ , one has $\mathbf{J}(\mathbf{x}') d^3x' = I d\mathbf{x}'$ where $d\mathbf{x}'$ is the vector tangent to the curve Γ .



In this case,

$$\mathbf{m} = I \frac{1}{2} \underbrace{\oint_{\Gamma} \mathbf{x}' \times d\mathbf{x}'}_{\mathbf{s}} = IS,$$

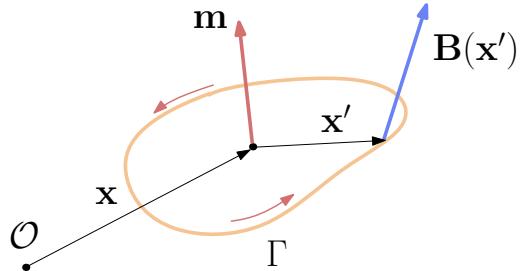
with $\mathbf{S} = S\mathbf{n}$ the oriented surface of the loop.

Remark

Contrary to what was obtained for an arbitrary charge distribution in electrostatics, here the first order approximation of the magnetic field generated by an arbitrary current is that of a magnetic dipole. In magnetostatics, there is no monopole term in the expansion.

7.8 Forces exerted on a magnetic dipole in an external magnetic field

Consider a magnetic dipole at position \mathbf{x} in the presence of an external magnetic field. In the same way as an external electric field is acting on an electric dipole, the external magnetic field will exert a force \mathbf{F} and a torque $\boldsymbol{\tau}$ on the magnetic dipole.



7.8.1 The Laplace force

A conductor carrying a current experiences a force \mathbf{F}_L , called the Laplace force, in the presence of an external magnetic field \mathbf{B} . It results from the Hall effect (see chapter 6), that is, the appearance of an electric field transverse to the velocity \mathbf{v} of the electrons and given by

$$\mathbf{E}_H = -\mathbf{v} \times \mathbf{B}.$$

While this Hall field cancels out the magnetic force acting on the electrons, it also exerts a force on the positive ions, which do not move. The electric force acting on a positive *ion* of charge $+e$ then reads

$$\mathbf{f} = e\mathbf{E}_H = -e\mathbf{v} \times \mathbf{B},$$

where \mathbf{v} is the velocity of the *electrons* in the conductor. If n is the density of ions (and electrons due to the global neutrality of the medium), the force acting on a differential volume d^3x' around \mathbf{x}' is

$$d\mathbf{f}(\mathbf{x}') = \mathbf{f}d^3x' = -env \times \mathbf{B}d^3x' = \mathbf{J} \times \mathbf{B}d^3x'.$$

For a wire conductor of section $\Delta S(\mathbf{x}')$, we can write $d^3x' = \Delta S(\mathbf{x}')dl(\mathbf{x}')$, where dl is a differential length element. The total force on d^3x' rewrites

$$d\mathbf{f}(\mathbf{x}') = \Delta S(\mathbf{x}') \underbrace{dl(\mathbf{x}')\mathbf{J}(\mathbf{x}') \times \mathbf{B}(\mathbf{x}')}_{dl(\mathbf{x}')J(\mathbf{x}')}$$

Defining $dl(\mathbf{x}')$ as the vector of magnitude $dl(\mathbf{x}')$ and whose direction coincides with that of the current density, and using the definition of the current $I = J\Delta S$, we finally have the Laplace force acting on a current element Idl ,

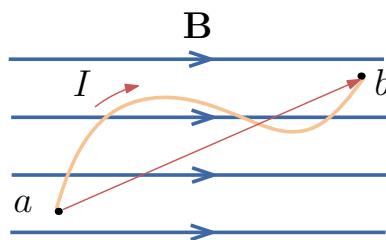
$$d\mathbf{F}_L(\mathbf{x}') = Idl(\mathbf{x}') \times \mathbf{B}(\mathbf{x}'). \quad (7.18)$$

The total Laplace force on a linear conductor described by the curve Γ is given by

$$\mathbf{F}_L = \int_{\Gamma} Idl \times \mathbf{B}(\mathbf{x}). \quad (7.19)$$

- **FORCE ON A CIRCUIT IN THE PRESENCE OF A UNIFORM FIELD**

As an example, consider a conductor that carries a current I in the presence of a uniform magnetic field \mathbf{B} , as shown in the figure



The force on this conductor is then given by

$$\mathbf{F} = I \left(\int_a^b d\mathbf{x} \right) \times \mathbf{B} = I (\mathbf{x}_b - \mathbf{x}_a) \times \mathbf{B}.$$

Let $\mathbf{l} = \mathbf{x}_b - \mathbf{x}_a$ be the vector directed from a to b . With this notation,

$$\mathbf{F} = I \mathbf{l} \times \mathbf{B}.$$

Now, if the conductor constitutes a closed circuit of arbitrary shape, the force will be

$$\mathbf{F} = I \left(\oint_{\Gamma} d\mathbf{x} \right) \times \mathbf{B}.$$

But

$$\oint_{\Gamma} d\mathbf{x} = \mathbf{0}.$$

The net force on a closed circuit in a uniform magnetic field is null, $\mathbf{F} = \mathbf{0}$.

Suggested exercises: Examples 7.18-7.20 in the exercise book.

7.8.2 Force on a magnetic dipole

To express this force, the dipole can be equivalently modeled as a closed circuit Γ carrying a current I or as a current distribution in a volume Ω .

Recall that the Laplace force that the field exerts on the current distribution is

$$\mathbf{f} = \iiint_{\Omega} \mathbf{j}(\mathbf{x}') \times \mathbf{B}(\mathbf{x}') d^3x'.$$

For a wire current, the corresponding expression reads

$$\mathbf{f} = \oint_{\Gamma} I d\mathbf{l}(\mathbf{x}') \times \mathbf{B}(\mathbf{x}').$$

If the magnetic field is uniform, then $\mathbf{B}(\mathbf{x}') = \mathbf{B}(\mathbf{x})$ and

$$\mathbf{f} = \underbrace{\oint_{\Gamma} I d\mathbf{x}' \times}_{\mathbf{0}} \mathbf{B}(\mathbf{x}) = \mathbf{0}.$$

If the magnetic field is nonuniform, consider each component of the force separately, for instance $f_x = \mathbf{f} \cdot \mathbf{u}_x$:

$$f_x = \mathbf{f} \cdot \mathbf{u}_x = \oint_{\Gamma} I d\mathbf{x}' \times \mathbf{B}(\mathbf{x}') \cdot \mathbf{u}_x = \oint_{\Gamma} I \mathbf{B}(\mathbf{x}') \times \mathbf{u}_x \cdot d\mathbf{x}'$$

where a circular permutation was used in the scalar triple product. Now the Stokes theorem allows us to transform the circulation of $\mathbf{B} \times \mathbf{u}_x$ into the flux of the curl of this vector through any surface $S(\Gamma)$ bounded by Γ

$$f_x = \iint_{S(\Gamma)} I \nabla' \times (\mathbf{B}(\mathbf{x}') \times \mathbf{u}_x) \cdot \mathbf{n}(\mathbf{x}') dS.$$

We use now a general identity of vectorial calculus: for any vector \mathbf{B} , we have $\nabla \times (\mathbf{B} \times \mathbf{u}_x) = \frac{\partial \mathbf{B}}{\partial x}$. Thus,

$$f_x = \iint_{S(\Gamma)} I \frac{\partial \mathbf{B}(\mathbf{x}')}{\partial x'} \cdot \mathbf{n}(\mathbf{x}') dS.$$

Finally, a Taylor expansion of $\frac{\partial \mathbf{B}(\mathbf{x}')}{\partial x'}$, considering $\mathbf{x}' = \mathbf{x} + (\mathbf{x}' - \mathbf{x})$, gives at leading order,

$$\frac{\partial \mathbf{B}(\mathbf{x}')}{\partial x'} \sim \frac{\partial \mathbf{B}(\mathbf{x})}{\partial x}.$$

We therefore recognize the dipole moment after factorizing the uniform term out of the integral, that is,

$$f_x = \iint_S I \frac{\partial \mathbf{B}(\mathbf{x})}{\partial x} \cdot \mathbf{n}(\mathbf{x}') dS = I \underbrace{\iint_{S(\Gamma)} \mathbf{n}(\mathbf{x}') dS}_{\mathbf{m}} \cdot \frac{\partial \mathbf{B}(\mathbf{x})}{\partial x}.$$

Finally, we obtain

$$f_x = \mathbf{m} \cdot \frac{\partial \mathbf{B}(\mathbf{x})}{\partial x}$$

(7.20)

and the components f_y and f_z are obtained by simply changing x into y or z .

Note that a nonuniform field is required in order to obtain a force. In the course of quantum physics, you will learn that Stern and Gerlach have discovered the quantization of angular momentum in atoms by exploiting precisely this magnetic force to deflect atoms in an inhomogeneous magnetic field.

Now if \mathbf{m} is a constant vector (it must not depend on \mathbf{B} to fulfill this condition), then ,

$$\mathbf{f} = -\nabla(-\mathbf{m} \cdot \mathbf{B}),$$

which allows us to identify an energy function for a permanent magnetic dipole \mathbf{m} in a magnetic field \mathbf{B}

$$W = -\mathbf{m} \cdot \mathbf{B}.$$

There is an obvious analogy with the expression for potential energy of an electric dipole in an electric field. However, there are subtle differences between forces acting on electric and magnetic dipoles.

Indeed, note that the expression $f_x = \mathbf{m} \cdot \frac{\partial \mathbf{B}(\mathbf{x})}{\partial x}$ was demonstrated from the expression for the Laplace force, which is valid even in the time-dependent regime. The force that a magnetic field exerts on a magnetic dipole is thus the same for time dependent fields $\mathbf{B}(\mathbf{x}, t)$.

Now for a stationary regime, the external magnetic field satisfies

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$$

where \mathbf{J} is here the current density that is the source of the external field and is located far away from the dipole. This means that $\mathbf{J}(\mathbf{x}) = \mathbf{0}$ and $\nabla \times \mathbf{B} = \mathbf{0}$. Starting from Eq. (7.20), we can then rewrite the force on the dipole as

$$f_x = m_x \frac{\partial B_x}{\partial x} + m_y \frac{\partial B_y}{\partial x} + m_z \frac{\partial B_z}{\partial x} = m_x \frac{\partial B_x}{\partial x} + m_y \frac{\partial B_x}{\partial y} + m_z \frac{\partial B_x}{\partial z}.$$

In other words, gathering all components, we find in the stationary regime,

$$f_x = (\mathbf{m} \cdot \nabla) B_x \quad \rightarrow \quad \mathbf{f} = (\mathbf{m} \cdot \nabla) \mathbf{B}.$$

This expression of the force is different from the general expression (7.20) obtained above. In the stationary regime, both expressions coincide but recall that only the general expression (7.20) must be used in the time-varying regime. A similar situation was encountered for the force on an electric dipole, however, remember that in a time-dependent electric field, the force exerted on an electric dipole was found to be $\mathbf{f} = (\mathbf{p} \cdot \nabla) \mathbf{E}$, that is, the general expression for the force applied to an electric dipole is analogous to the particular expression for the magnetic dipole in the stationary regime, and vice versa.

7.8.3 Torque on a magnetic dipole

The torque applied on the magnetic dipole in the presence of the magnetic field is obtained in a similar way as the force, by integration of the elementary torques resulting from the Laplace forces applied to the contour that model the magnetic dipole:

$$\boldsymbol{\tau}(\mathbf{x}) = \oint_{\Gamma} \mathbf{x}' \times (i d\mathbf{x}' \times \mathbf{B}(\mathbf{x}')).$$

It is sufficient to consider now the Taylor expansion of the magnetic field

$$\mathbf{B}(\mathbf{x}') = \mathbf{B}(\mathbf{x}) + [(\mathbf{x}' - \mathbf{x}) \cdot \nabla] \mathbf{B}(\mathbf{x}) + \dots,$$

at leading order, that is, $\mathbf{B}(\mathbf{x}') = \mathbf{B}(\mathbf{x})$, so as to get the torque exerted on the dipole,

$$\begin{aligned} \boldsymbol{\tau}(\mathbf{x}) &= \oint_{\Gamma} \mathbf{x}' \times [i d\mathbf{x}' \times \mathbf{B}(\mathbf{x})] \\ &= \oint_{\Gamma} i[\mathbf{x}' \cdot \mathbf{B}(\mathbf{x})] d\mathbf{x}' - \underbrace{\oint_{\Gamma} i(\mathbf{x}' \cdot d\mathbf{x}') \mathbf{B}(\mathbf{x})}_0 \end{aligned}$$

The second term on the right-hand side is zero because $\oint_{\Gamma} (\mathbf{x}' \cdot d\mathbf{x}') = \frac{1}{2} \oint_{\Gamma} d\mathbf{x}'^2 = 0$. Then, consider separately the components for the torque

$$\tau_x = \boldsymbol{\tau} \cdot \mathbf{u}_x = \int_{\Gamma} i[\mathbf{x}' \cdot \mathbf{B}(\mathbf{x})] \mathbf{u}_x \cdot d\mathbf{x}'$$

The Stokes theorem gives

$$\tau_x = \iint_S i \underbrace{\nabla' \times [(\mathbf{x}' \cdot \mathbf{B}(\mathbf{x})) \mathbf{u}_x]}_{\mathbf{B}(\mathbf{x}) \times \mathbf{u}_x} \cdot \mathbf{n}(\mathbf{x}') dS$$

Then,

$$\tau_x = \iint_S i \mathbf{B}(\mathbf{x}) \times \mathbf{u}_x \cdot \mathbf{n} dS = \underbrace{\left[\iint_S i \mathbf{n}(\mathbf{x}') dS \right]}_{\mathbf{m}} \times \mathbf{B}(\mathbf{x}) \cdot \mathbf{u}_x$$

Finally,

$$\boldsymbol{\tau}(\mathbf{x}) = \mathbf{m} \times \mathbf{B}(\mathbf{x}).$$

7.9 The two fundamental laws of magnetostatics

Below we will enunciate the two fundamental laws of magnetostatics. Firstly, we will see that it is possible to define a vector potential \mathbf{A} such that the magnetic field can be obtained as the curl of the potential, $\mathbf{B} = \nabla \times \mathbf{A}$. In consequence, the divergence of the magnetic field is always zero, $\nabla \cdot \mathbf{B} = 0$. The latter equation, together with the differential form of Ampère's law, form a complete set of equations that determine uniquely the magnetic field at every point in space (Helmholtz's decomposition theorem).

7.9.1 Gauss's law for magnetism

The divergence of a curl is always zero, that is

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0,$$

so that the differential equation for the divergence of \mathbf{B} is simply

$$\nabla \cdot \mathbf{B} = 0. \tag{7.21}$$

The integral form of Eq. (7.21) can be obtained from Green-Ostrogradsky's theorem:

$$\iint_S \mathbf{B}(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}') = \iiint_{\Omega(S)} \nabla \cdot \mathbf{B}(\mathbf{x}') d^3 x' = 0, \tag{7.22}$$

so that the magnetic flux over any closed surface is always zero. As will be shown later, this fundamental equation remains valid even in the case of distributions and fields depending on time. This equation means that magnetic monopoles (from which magnetic field lines would start or end) do not exist.

7.9.2 The two fundamental laws of magnetostatics

So far, the fundamental principles of magnetostatics have been stated as being Biot-Savart's law for a moving point charge and the superposition principle. From these principles, it has been possible to demonstrate the two following laws:

1. Gauss's law for magnetism: the magnetic flux over any closed surface is zero or, equivalently, the divergence of the magnetic field is zero:

$$\iint_{\partial\Omega} \mathbf{B}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}) = 0 \Leftrightarrow \nabla \cdot \mathbf{B} = 0.$$

2. Ampère's law: the circulation of the magnetic field over any closed curve Γ is proportional to the current passing through any surface $S(\Gamma)$ bounded by the curve Γ

$$\oint_{\Gamma} \mathbf{B}(\mathbf{x}) \cdot d\mathbf{x} = \mu_0 I_{\Gamma} \Leftrightarrow \nabla \times \mathbf{B} = \mu_0 \mathbf{J}.$$

The differential forms of these two laws correspond to the two Maxwell equations for static magnetic fields. From the Helmholtz theorem, it is seen that magnetostatics is a complete theory.

7.10 Completeness of magnetostatics

There is a unique field that decays faster than $1/|\mathbf{x}|^2$ at infinity that verifies:

$$\begin{cases} \nabla \cdot \mathbf{B} = 0, \\ \nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \end{cases}$$

Indeed, Helmholtz's theorem states that \mathbf{B} is completely determined by the following integral-differential representation:

$$\mathbf{B}(\mathbf{x}) = \nabla \times \frac{1}{4\pi} \iiint_{\mathbb{R}^3} \frac{\nabla' \times \mathbf{B}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' = \nabla \times \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'.$$

Using the identities $\nabla \times (f \mathbf{J}(\mathbf{x}')) = f \underbrace{\nabla \times \mathbf{J}(\mathbf{x}')}_0 + \nabla f \times \mathbf{J}(\mathbf{x}')$ where the scalar $f = 1/|\mathbf{x} - \mathbf{x}'|$, we find

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x', \quad (7.23)$$

which is Biot-Savart's law. The set of the two fundamental differential equations is therefore sufficient to determine the magnetic field.

7.11 Summary and essential formulas

- Moving electrical charges generate a magnetic field \mathbf{B} , which is a vector field defined over the entire space $\mathbf{B} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. There is a deep relationship between electric and magnetic fields. Indeed, considering a reference system S' in which a charge is at rest, an observer will see only an electrostatic field, while in any other reference system, an observer will see a charge in motion, and then, a magnetic field. Electricity and magnetism are aspects of the same phenomenon.
- The magnetic field generated by a current density $\mathbf{J} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is given by the integral equation (7.7)

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} d^3x' \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}.$$

- The field generated by a linear current defined by the curve Γ is given by the Biot-Savart law, which is a particular case of Eq. (7.7):

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 I}{4\pi} \int_{\Gamma} d\mathbf{l}(\mathbf{x}') \times \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}.$$

- Thanks to the Dirac distribution, the current density associated with a point charge q , of speed \mathbf{v} and located at \mathbf{x}_0 can be written as a volume density $\mathbf{J}(\mathbf{x}) = q\mathbf{v}\delta(\mathbf{x} - \mathbf{x}_0)$. In this way, the magnetic field of a point charge is obtained:

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} d^3x' \frac{q\mathbf{v}\delta(x - x_0) \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = \frac{\mu_0 q \mathbf{v}}{4\pi} \times \frac{(\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^3}.$$

- The force experienced by a charge q in the presence of a magnetic field \mathbf{B} is

$$\mathbf{F}_q = q\mathbf{v} \times \mathbf{B}.$$

This force is responsible for a wide variety of phenomena. In a conductive medium carrying a current I , the application of a magnetic field \mathbf{B} perpendicular to the current generates a voltage difference in the direction transverse to the current, called Hall voltage and given by $V_H = -\frac{BI}{qn b}$, where b is the thickness of the conductor and n the density of carriers. The Hall effect can be used to measure magnetic fields, by applying a current through a conductor of known conductivity, and by measuring V_H , it is possible to determine B .

- A magnetic moment \mathbf{m} can be associated with a closed current I . If the current is described by a plane curve that encloses a surface S of normal \mathbf{n} , we have

$$\mathbf{m} = IS\mathbf{n}.$$

- The magnetic field generated by a current density $\mathbf{J} : \Omega \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is given by Biot-Savart's integral (7.8):

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'.$$

- Ampère's law is a consequence of Biot-Savart's law (7.8) and states that the circulation of \mathbf{B} along any closed curve Γ is proportional to the current I_Γ flowing through any surface $S(\Gamma)$ whose bounding curve is Γ :

$$\oint_{\Gamma} d\mathbf{x} \cdot \mathbf{B}(\mathbf{x}) = \mu_0 I_\Gamma.$$

In its differential form, Ampère's law can be written as

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}.$$

- The magnetic field derives from a vector potential \mathbf{A} such that $\mathbf{B} = \nabla \times \mathbf{A}$. This implies a null divergence for the magnetic field

$$\nabla \cdot \mathbf{B} = 0.$$

Equivalently, the magnetic flux over any closed surface is always zero,

$$\iint_S d\mathbf{S} \cdot \mathbf{B} = 0.$$

The vector potential is not uniquely defined, unless one adds a constraint such as $\nabla \cdot \mathbf{A} = 0$ (Coulomb's gauge), for which \mathbf{A} is uniquely defined by

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x',$$

which corresponds to the solution of Poisson's equation

$$\nabla^2 \mathbf{A}(\mathbf{x}) = -\mu_0 \mathbf{J}(\mathbf{x}).$$

- Far from an arbitrary current distribution, the first order approximation for the vector potential is that of a magnetic dipole

$$\mathbf{A}(\mathbf{r}) \approx \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{r}}{r^3}$$

and the magnetic field has a similar expression as the electric field generated by an electric dipole

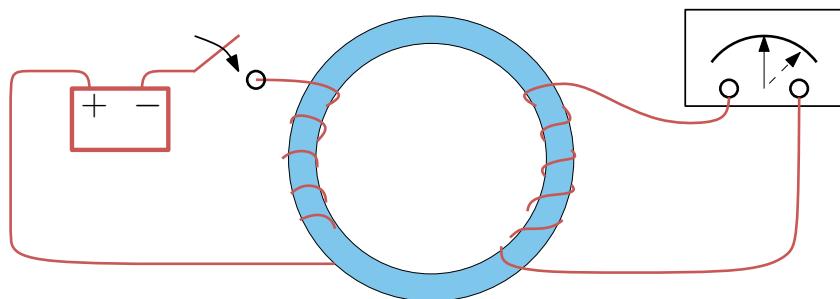
$$\mathbf{B}(\mathbf{r}) = -\nabla \left(\frac{\mu_0}{4\pi} \frac{\mathbf{m} \cdot \mathbf{r}}{r^3} \right).$$

Chapter 8

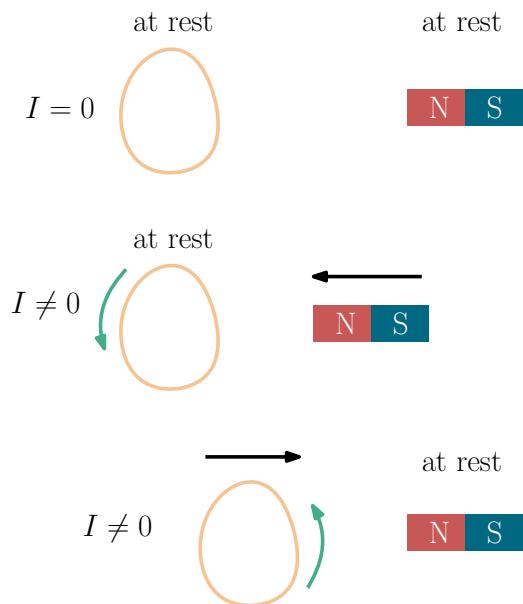
Faraday's law of induction, magnetic forces and energy

Introduction

The physicists Oersted, Ampère, Biot and Savart established a first relationship between electricity and magnetism: an electric field in a conductor produces an electric current and therefore a magnetic field. Intuition led them to anticipate that the inverse process may exist as well, that is, a magnetic field should be capable of producing electricity. One decade after the discoveries of Oersted, Michael Faraday discovered in 1831 the phenomenon of mutual inductance, in which he used two isolated coils wound on an iron ring, and observed that when a current through one of the coils suddenly changes, a current is induced momentarily in the second coil.



In a second series of experiments, Faraday observed that a current was induced in a conductor loop whenever a magnet was moved with respect to it. He also observed that the same phenomenon occurred when the magnet was fixed and the loop was moving with respect to the magnet. These phenomena are examples of *electromagnetic induction*.

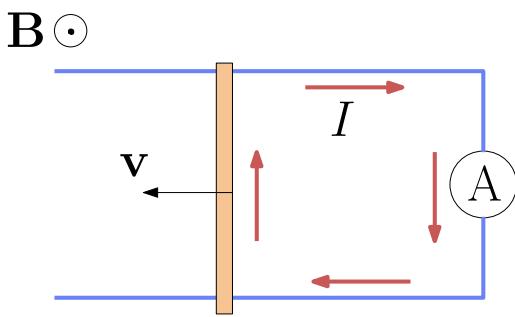


Faraday's phenomenon of induction is nowadays used all around the globe to produce electricity. The first power plant in history was inaugurated in London in 1882, in which enormous steam turbines made coils turn around a magnet. According to Faraday's law of induction, a current is generated in the coil and thermal energy is thus transformed into electric energy. Almost every power plant works according to this principle, the energy to turn the coils may differ from one type of plant to the other: it can be extracted from various source such as the flux of water in a river, the wind, or the pressure of water heated by nuclear reactions. Conversely, electric currents can produce a mechanical rotation and that is how motors work. Faraday's law of induction is at the heart of the electromechanical conversion of energy.

8.1 Induced currents: experimental observations

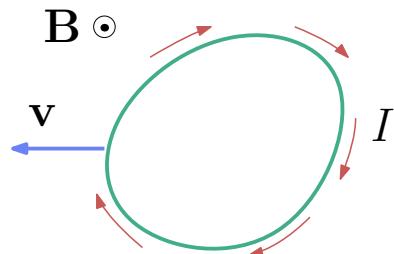
8.1.1 Deformation of a circuit in a static magnetic field

Consider a region in which there is a static magnetic field \mathbf{B} perpendicular to a closed circuit, as shown in the figure below. A conductive bar closes the circuit and can move freely. When the bar is displaced, an induced current I is generated in the circuit. If the direction of the displacement is reversed, so does the direction of the induced current. In the absence of a magnetic field, no current is observed. In conclusion, an induced current appears when a displacement of a part of the circuit occurs in the presence of a static magnetic field.



8.1.2 Mobile circuit in a static magnetic field

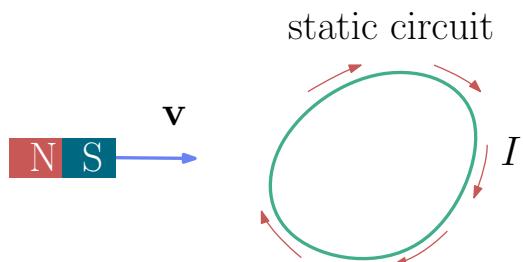
The same phenomenon occurs when the whole circuit is in relative motion with respect to a static, non-uniform magnetic field: an induced current is generated in the circuit.



The reason why the magnetic field has to be non-uniform is that no current is generated if the flux of the field through the circuit does not change during the motion of the circuit (A relative motion in a uniform field does not change the flux). It will become clear below.

8.1.3 Circuit at rest in a time-varying magnetic field

This is the case obtained when the device creating the magnetic field is displaced. For example, this can be a magnet moving relative to a static circuit.



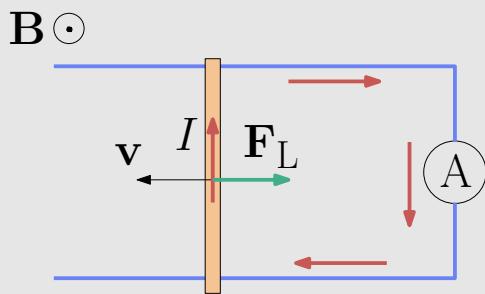
Of course, the phenomenon of induction depends solely on the relative motion between the loop and the magnet. The case of a mobile circuit in a static magnetic field and the case of a time-varying magnetic field in a static circuit both describe the exact same physical phenomenon, they are related to each other by a change of reference frame.

8.1.4 Lenz's law (1834)

All the previous experimental observations led to the conclusion that the magnitude of the induced current is directly proportional to the variation of the magnetic flux through the circuit. In addition, Emil Lenz stated that the direction of the induced current is such that it opposes the variation of the magnetic flux through the circuit.

Example 8.1 - Laplace rail

A conductive bar on top of a rectangular circuit is displaced with velocity \mathbf{v} in the presence of a uniform magnetic field \mathbf{B} perpendicular to the plane of the circuit. According to Lenz's law, the induced current flows in the direction indicated in the figure below. The surface of the circuit is increasing, leading to an increase of the magnetic flux through the circuit. The Laplace force \mathbf{F}_L due to the induced current must act on the bar against the change of magnetic flux in the circuit and the original displacement, thus the induced current must flow clockwise.

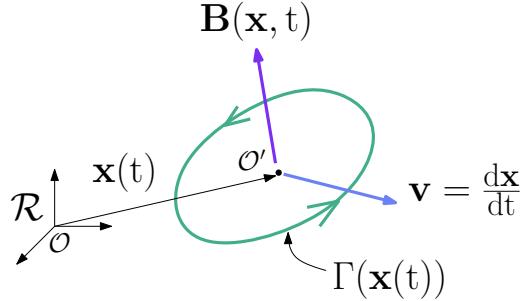


8.2 Electromotive force - Faraday's law

In this section we will establish the induction law which unifies all the phenomena discovered by Faraday as well as Lenz's law. In the case of a mobile circuit in a static magnetic field, the induction law is simply a consequence of the Lorentz force acting on the electrons of the moving circuit. In contrast, in the reference frame in which the circuit is at rest and the magnetic field varies in time, this same phenomenon cannot be explained by means of the Lorentz force, since the latter is zero in this reference frame. We will see that it is then necessary to admit the appearance of an electric field generated by the varying magnetic field which is responsible for the movement of charges in the circuit.

8.2.1 Magnetic flux through a moving surface

Consider a closed curve Γ centered around a point \mathbf{x} . The curve moves at velocity \mathbf{v} in a reference frame \mathcal{R} in the presence of a time-dependent magnetic field $\mathbf{B} = \mathbf{B}(\mathbf{x}, t)$.



If $S(\Gamma)$ is any surface whose bounding curve is Γ , the magnetic flux through $S(\Gamma)$ is defined as

$$\Phi_{S(\Gamma), \mathbf{B}} = \iint_{S(\Gamma)} \mathbf{B}(\mathbf{x} + \mathbf{x}', t) \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}'),$$

where $\mathbf{n}(\mathbf{x}')$ denotes the normal to the surface at $\mathbf{x}' \in S(\Gamma)$. An orientation of the curve Γ and a corresponding orientation of the normal to the surface are chosen, which amounts to defining one side of the surface as the positive side¹. The normal vector points toward the positive side and the orientation of Γ is anticlockwise, if viewed from the positive side. The SI-unit of the magnetic flux is the Weber (Wb):

$$1 \text{ Wb} = 1 \text{ T m}^2.$$

As a consequence of Maxwell-Thomson's equation $\nabla \cdot \mathbf{B} = 0$, the flux $\Phi_{S(\Gamma), \mathbf{B}}$ does not depend on the choice of the surface $S(\Gamma)$, as long as Γ is bounding it². $\Phi_{S(\Gamma), \mathbf{B}}$ will, in general, depend on the position \mathbf{x} of the loop, but also on time if the magnetic field is not static, that is $\Phi_{S(\Gamma), \mathbf{B}} = \Phi_{S(\Gamma), \mathbf{B}}(\mathbf{x}(t), t)$. The total derivative of Φ with respect to time, applying the chain rule, is then

$$\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dt} = \frac{\partial}{\partial t} \Phi_{S(\Gamma), \mathbf{B}} + \underbrace{\frac{d\mathbf{x}}{dt}}_{\mathbf{v}} \cdot \nabla \Phi_{S(\Gamma), \mathbf{B}}.$$

The derivative of the loop position with respect to time is simply the loop velocity. The last term $(\mathbf{v} \cdot \nabla) \Phi_{S(\Gamma), \mathbf{B}}$, called *advection* term, then represents the change in magnetic flux that is due to the motion of the loop. Finally, the total derivative with respect to time of the magnetic flux is:

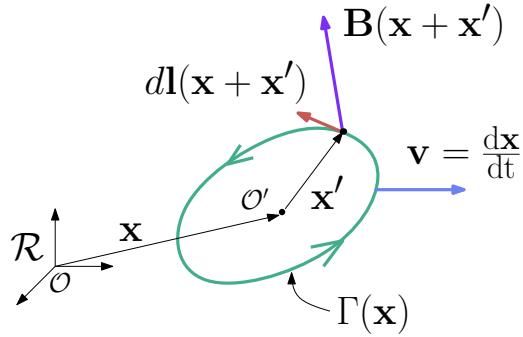
$$\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dt} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \Phi_{S(\Gamma), \mathbf{B}}. \quad (8.1)$$

¹We consider only orientable surfaces, i.e., surfaces that admit a continuous unit normal vector over the entire surface. This is not the case for exotic surfaces like the Moebius strip, therefore out of scope of the present course.

²Make sure you know how to prove this assertion: Hint: consider the two pieces of a closed surface bounded by Γ and apply Ostrogradsky's theorem.

8.2.2 Mobile circuit in a static magnetic field

Consider a closed, rigid circuit Γ whose center O' moves at velocity $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ with respect to the *laboratory* reference frame \mathcal{R} , in the presence of a static magnetic field \mathbf{B} . In this case, an induced current around Γ can be easily explained by means of the Lorentz force.



At a given instant, a charge element δq located at \mathbf{x}' with respect to O' moves with a velocity $\mathbf{w} = \mathbf{v} + \mathbf{u}$, where \mathbf{u} is the relative velocity of the charges with respect to the circuit, i.e., parallel to $d\mathbf{l}(\mathbf{x}')$. The magnetic component of the Lorentz force acting on δq is given by

$$d\mathbf{F}(\mathbf{x}') = \delta q (\mathbf{w} \times \mathbf{B})$$

and does not work since the magnetic Lorentz force is perpendicular to the trajectory of the charge

$$0 = \delta q (\mathbf{w} \times \mathbf{B}) \cdot \mathbf{w} dt.$$

Writing $\mathbf{w} = \mathbf{v} + \mathbf{u}$, where \mathbf{u} denotes the velocity of electrons with respect to the conductor, we can expand the scalar triple product and keep the only two non-zero terms. We still obtain a zero work:

$$0 = \delta q (\mathbf{v} \times \mathbf{B}) \cdot \mathbf{u} dt + \delta q (\mathbf{u} \times \mathbf{B}) \cdot \mathbf{v} dt.$$

In the second term, we recognize the Hall field $\mathbf{E}_H = -\mathbf{u} \times \mathbf{B}$ that is at the origin of Laplace's force, as commented below. Now, $d\mathbf{l}(\mathbf{x}') = \mathbf{u} dt$ and integration at a given instant over the contour Γ yields

$$0 = \delta q \underbrace{\oint_{\Gamma(\mathbf{x})} (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}(\mathbf{x}')}_{\varepsilon} + \frac{\delta q}{\delta t} \oint_{\Gamma(\mathbf{x})} (d\mathbf{l}(\mathbf{x}') \times \mathbf{B}) \cdot \mathbf{v} \delta t. \quad (8.2)$$

The first term can be identified as the product of the charge δq by the quantity

$$\varepsilon = \oint_{\Gamma(\mathbf{x})} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l}(\mathbf{x}')$$

which represents the circulation around the loop of a field $\mathbf{E}_{\text{emf}} = \mathbf{v} \times \mathbf{B}$, called *electromotive field*. The electromotive field is an electric field and has the dimension of a force per unit

charge. It is important to note that \mathbf{v} is not the velocity of the electrons but the velocity of the circuit, that is, the velocity of the fixed ions of the conductor. For a differential element of displacement $\delta\mathbf{x} = \mathbf{v}\delta t$ of the loop, the second term in Eq. (8.2) represents the work $\mathbf{F}_L \cdot \mathbf{v}\delta t$ of the net Laplace force $\mathbf{F}_L = i \oint_{\Gamma(\mathbf{x})} (d\mathbf{l}(\mathbf{x}') \times \mathbf{B})$ acting on the loop, where $i = \delta q/\delta t$ denotes the current flowing in the loop. Equation (8.2) can now be interpreted: To move the loop at velocity \mathbf{v} , an operator exerts a force that is opposite to the net Laplace force \mathbf{F}_L , and thus produce a work that is equal to $\delta q \varepsilon$. Thus, the operator moves the circuit by providing a work per unit charge, ε , called the induced electromotive force (emf). However, ε is not a force as its name would suggest since it has dimensions of a voltage or a work per unit charge.

The electromotive field $\mathbf{E}_{\text{emf}} = \mathbf{v} \times \mathbf{B}$ here established by the magnetic force is responsible for the movement of electrons forming the induced current in the circuit. Note that in the reference frame, \mathbf{E}_{emf} is identified as an electric field but it represents a pure magnetostatic force per unit charge. By the Stokes theorem, the electromotive force then writes

$$\varepsilon = \iint_{S(\Gamma(\mathbf{x}))} \nabla' \times (\mathbf{v} \times \mathbf{B}(\mathbf{x} + \mathbf{x}')) \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}') = - \iint_{S(\Gamma(\mathbf{x}))} (\mathbf{v} \cdot \nabla') \mathbf{B}(\mathbf{x} + \mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}')$$

where we have used the identity

$$\nabla' \times (\mathbf{v} \times \mathbf{B}) = \mathbf{v} \underbrace{(\nabla' \cdot \mathbf{B})}_0 - \mathbf{B} \cdot \underbrace{(\nabla' \cdot \mathbf{v})}_0 + \underbrace{(\mathbf{B} \cdot \nabla') \mathbf{v}}_0 - (\mathbf{v} \cdot \nabla') \mathbf{B}.$$

Finally, since $(\mathbf{v} \cdot \nabla') \mathbf{B}(\mathbf{x} + \mathbf{x}') = (\mathbf{v} \cdot \nabla) \mathbf{B}(\mathbf{x} + \mathbf{x}')$,

$$\varepsilon = -(\mathbf{v} \cdot \nabla) \underbrace{\iint_{\Gamma(\mathbf{x})} \mathbf{B}(\mathbf{x} + \mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}')}_{\Phi_{S(\Gamma(\mathbf{x}))}, \mathbf{B}}$$

and

$$\varepsilon = \oint_{\Gamma(\mathbf{x})} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -(\mathbf{v} \cdot \nabla) \Phi_{S(\Gamma(\mathbf{x}))}, \mathbf{B}. \quad (8.3)$$

We recognize on the right-hand side of Eq. (8.3) the advection term of the time derivative of the magnetic flux (8.1). Since the magnetic field is static, it corresponds as well to the total time derivative of the magnetic field. Finally,

$$\varepsilon = \oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -\frac{d}{dt} \Phi_{S(\Gamma)}, \mathbf{B}. \quad (8.4)$$

Since the flux has units of Wb, we see that ε has units of $\text{Wb s}^{-1} = \text{V}$. Note that the minus sign in Eq. (8.4) is consistent with Lenz's law and indicates that the induced current on the loop opposes to the change of magnetic flux through the circuit. Note also that Eq. (8.4) was found by using solely the laws of magnetostatics.

Remarks

- In the reference frame \mathcal{R}' in which the circuit is at rest, i.e., \mathcal{R}' moves at \mathbf{v} with respect to \mathcal{R} , the Lorentz force acting on the electrons should be the same as in the laboratory frame. Since in \mathcal{R}' the circuit is at rest, this is only possible if in \mathcal{R}' the electrons are subjected to an electric field \mathbf{E}' such that

$$\mathbf{E}' = \mathbf{v} \times \mathbf{B}.$$

More generally, if \mathbf{E} and \mathbf{B} are, respectively, the electric and magnetic fields in the reference frame \mathcal{R} , the fields in the reference frame \mathcal{R}' are given, in the non-relativistic limit $|\mathbf{v}| \ll c$, by

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \quad \mathbf{B}' = \mathbf{B}.$$

This transformation leaves the Lorentz force invariant in all reference frames. This remarkable result shows that the electric or magnetic character of the electromotive field $\mathbf{E}_{\text{emf}} = \mathbf{v} \times \mathbf{B}$ depends on the reference frame.

8.2.3 Static circuit in a time-varying magnetic field

In this case, the laws of magnetostatics alone are not capable of explaining the experimental results. Nevertheless, it was experimentally shown that an electromotive field, and thus an electromotive force responsible for the appearance of an induced current still satisfies

$$\varepsilon = -\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dt},$$

where the variation of the magnetic flux comes, in this case, from a temporal variation of the magnetic field itself and not from a movement of the loop Γ . We then see that a time-varying magnetic field \mathbf{B} generates an electromotive field, that is a circulating electric field \mathbf{E}_{emf} that puts the electrons in movement in the current loop. In other words, the electromotive field \mathbf{E}_{emf} in this case coincides with the electric field generated by the time-varying magnetic field \mathbf{B} and we have

$$\varepsilon = \oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -\frac{d}{dt} \iint_{S(\Gamma)} \mathbf{B} \cdot d\mathbf{S}.$$

Now, using the fact that \mathbf{B} derives from a vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$ and by using the Stokes theorem, the emf reads

$$\oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -\frac{d}{dt} \iint_{S(\Gamma)} (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = -\frac{d}{dt} \oint_{\Gamma} \mathbf{A} \cdot d\mathbf{l}.$$

Finally,

$$\oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -\oint_{\Gamma} \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l}$$

and the electromotive field writes

$$\mathbf{E}_{\text{emf}} = -\frac{\partial \mathbf{A}}{\partial t}.$$

Definition 8.1: Faraday's law of induction (1831)

In the most general case of a circuit Γ moving at velocity \mathbf{v} in the presence of a time-varying magnetic field \mathbf{B} , an electromotive field \mathbf{E}_{emf} produces a current in the circuit. This field is given by

$$\mathbf{E}_{\text{emf}} = \mathbf{v} \times \mathbf{B} - \frac{\partial \mathbf{A}}{\partial t}. \quad (8.5)$$

One recognizes in the first term the component of the electromotive field caused by the motion of the circuit in the magnetic field, whereas the second term corresponds to the electric field generated by a time-dependent magnetic field. The electromotive force around a closed loop Γ is the circulation of \mathbf{E}_{emf} around the loop, and Faraday's law states that the electromotive force is given by the total derivative with respect to time of the magnetic flux through the circuit

$$\varepsilon = \oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = - \left(\mathbf{v} \cdot \nabla + \frac{\partial}{\partial t} \right) \iint_{S(\Gamma)} \mathbf{B} \cdot d\mathbf{S} = - \frac{d}{dt} \iint_{S(\Gamma)} \mathbf{B} \cdot d\mathbf{S}$$

so that, in the most general case

$$\varepsilon = - \frac{d}{dt} \iint_{S(\Gamma)} \mathbf{B} \cdot d\mathbf{S} = - \frac{d}{dt} \Phi_{S(\Gamma), \mathbf{B}}. \quad (8.6)$$

8.3 Faraday's law, differential form

As we have seen in electrostatics, a distributions of charges constitutes a source for a curl free electric field, \mathbf{E}_{es} , satisfying $\nabla \times \mathbf{E}_{\text{es}} = \mathbf{0}$. Faraday's law of induction states that a time-varying magnetic flux is the source for an electromotive field. Adding both fields, the electric field is

$$\mathbf{E} = \mathbf{E}_{\text{es}} + \mathbf{E}_{\text{emf}}.$$

Let Γ be an arbitrary, static closed curve (not necessarily representing a physical current loop). The circulation of the electric field \mathbf{E} along Γ can be rewritten, thanks to the Stokes theorem, as:

$$\iint_{S(\Gamma)} (\nabla \times \mathbf{E}(\mathbf{x}, t)) \cdot d\mathbf{S}(\mathbf{x}) = \oint_{\Gamma} \mathbf{E}(\mathbf{x}, t) \cdot d\mathbf{l}(\mathbf{x}).$$

On the other hand,

$$\oint_{\Gamma} \mathbf{E}(\mathbf{x}, t) \cdot d\mathbf{l}(\mathbf{x}) = \oint_{\Gamma} \mathbf{E}_{\text{es}}(\mathbf{x}, t) \cdot d\mathbf{l}(\mathbf{x}) + \oint_{\Gamma} \mathbf{E}_{\text{emf}}(\mathbf{x}, t) \cdot d\mathbf{l}(\mathbf{x}) = \varepsilon,$$

owing to the fact that

$$\oint_{\Gamma} \mathbf{E}_{\text{es}}(\mathbf{x}, t) \cdot d\mathbf{l}(\mathbf{x}) = \iint_{S(\Gamma)} \underbrace{(\nabla \times \mathbf{E}_{\text{es}}(\mathbf{x}, t))}_{\mathbf{0}} \cdot d\mathbf{S}(\mathbf{x}) = \mathbf{0}.$$

Now Faraday's law of induction states that

$$\varepsilon = -\frac{d}{dt} \iint_{S(\Gamma)} \mathbf{B} \cdot d\mathbf{S} = -\iint_{S(\Gamma)} \frac{\partial}{\partial t} \mathbf{B}(\mathbf{x}, t) \cdot d\mathbf{S}(\mathbf{x})$$

Then

$$\iint_{S(\Gamma)} \left(\nabla \times \mathbf{E}(\mathbf{x}, t) + \frac{\partial}{\partial t} \mathbf{B}(\mathbf{x}, t) \right) \cdot d\mathbf{S}(\mathbf{x}) = 0,$$

and since Γ is an arbitrary curve, we obtain the differential form of Faraday's law of induction

$$\boxed{\nabla \times \mathbf{E}(\mathbf{x}, t) = -\frac{\partial}{\partial t} \mathbf{B}(\mathbf{x}, t).} \quad (8.7)$$

This is one of the four Maxwell equations for time-varying electromagnetic fields and it is a generalization of the law of electrostatics $\nabla \times \mathbf{E} = \mathbf{0}$. A remarkable consequence is that the electric field is no longer conservative in the time-varying regime. Using the vector potential in Eq. (8.7), we obtain

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t} (\nabla \times \mathbf{A}),$$

where the order or time and space derivatives can be exchanged:

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0.$$

It follows that it is the field $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$, which differs from the electric field \mathbf{E} , that is conservative and derives from a potential V . One concludes that

$$\boxed{\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t},} \quad (8.8)$$

where V is the usual electrostatic potential in Coulomb's gauge ($\nabla \cdot \mathbf{A} = 0$). Indeed, $\nabla \cdot \mathbf{E} = -\nabla^2 V = \frac{\rho}{\epsilon_0}$. The first term on the right-hand side of Eq. (8.8) represents therefore a conservative electric field generated by charge separation ($\rho \neq 0$), whereas the second term is a non-conservative component ($\nabla \times (\partial_t \mathbf{A}) \neq 0$) due to the existence of a time-varying magnetic field.

8.3.1 Generalized Ohm's law

Consider a conducting wire moving at velocity \mathbf{v} in an external magnetic field \mathbf{B} . In addition to the electric field \mathbf{E} , the electrons in the wire will feel a force $-e(\mathbf{v} + \mathbf{u}) \times \mathbf{B}$ due to the magnetic component of the Lorentz force, with \mathbf{u} the velocity of electrons with respect to the wire. The equation of motion writes

$$m_e \frac{d}{dt}(\mathbf{v} + \mathbf{u}) + \frac{m_e}{\tau} \mathbf{u} = -e\mathbf{E} - e(\mathbf{v} + \mathbf{u}) \times \mathbf{B}.$$

In the steady-current approximation, we will neglect the acceleration term, so that at a given time t , we apply the laws of the steady regime. This is a good approximation as long as the fields vary at frequencies lower than 10 GHz. We then have:

$$\mathbf{u} = -\frac{e\tau}{m_e} (\mathbf{E} + (\mathbf{v} + \mathbf{u}) \times \mathbf{B}).$$

The current density in the wire is $\mathbf{J} = -en\mathbf{u}$ and therefore

$$\mathbf{J} = \underbrace{\frac{ne^2\tau}{m_e}}_{\sigma} (\mathbf{E} + (\mathbf{v} + \mathbf{u}) \times \mathbf{B}).$$

This equation can be projected on the directions parallel and perpendicular to \mathbf{J} . By definition, \mathbf{J} is parallel to \mathbf{u} and has no perpendicular component to itself. This means that the perpendicular projection yields

$$\underbrace{\mathbf{E}_\perp + (\mathbf{v} \times \mathbf{B})_\perp}_{\mathbf{E}'_\perp} = \underbrace{-\mathbf{u} \times \mathbf{B}}_{\mathbf{E}_{\text{Hall}}}.$$

The left-hand side represents the perpendicular component of the electric field \mathbf{E}'_\perp exerted on electrons in the frame of the conductor and is equal to the Hall field.

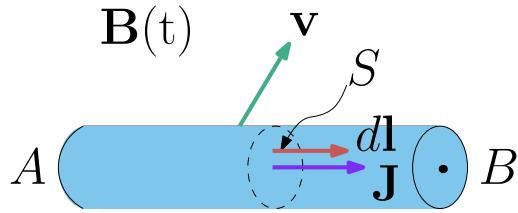
We obtain a generalized Ohm's law from the parallel projection:

$$\mathbf{J} = \sigma (\mathbf{E} + \mathbf{v} \times \mathbf{B})_{\parallel}$$

and, according to Eq. (8.8), in the presence of a time-varying magnetic field, the electric field writes $\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}$ and so:

$$\mathbf{J} = \sigma \left(-\nabla V - \underbrace{\frac{\partial \mathbf{A}}{\partial t}}_{\mathbf{E}_{\text{emf}}} + \mathbf{v} \times \mathbf{B} \right)_{\parallel}$$

where we recognize the electromotive field \mathbf{E}_{emf} . Supposing that the wire has a section S , let us integrate the generalized Ohm's law between two points A and B in the conductor along the current field lines ($d\mathbf{l} \parallel \mathbf{J}$):



$$\int_A^B \mathbf{J} \cdot d\mathbf{l} = -\sigma \int_A^B \nabla V \cdot d\mathbf{l} + \sigma \int_A^B \left(-\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B} \right) \cdot d\mathbf{l}.$$

Now, the current density can be written as $\mathbf{J} = \frac{I}{S} \frac{d\mathbf{l}}{dl}$ where I is the current flowing from A to B . We then have $\mathbf{J} \cdot d\mathbf{l} = \frac{dl}{S}$ and so:

$$I \underbrace{\int_A^B \frac{dl}{\sigma S}}_{R_{AB}} = V_A - V_B + \underbrace{\int_A^B \left(-\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B} \right) \cdot d\mathbf{l}}_{\varepsilon_{AB}}.$$

We recognize in R_{AB} the resistance between A and B , and ε_{AB} is the path integral of the electromotive field between A and B . Finally the integral form of Ohm's law writes

$$V_A - V_B = R_{AB}I - \varepsilon_{AB}. \quad (8.9)$$

Remarks

- Since the electromotive field is not conservative, the potential difference $V_A - V_B$ will, in general, depend on the path that is taken from A to B .
- For an open circuit ($I = 0$), we obtain

$$V_B - V_A = \varepsilon_{AB}.$$

The induction force pushes positive charges on one side, negative charges on the other side, thus creating a charge separation responsible of the appearance of a potential difference $V_B - V_A$.

- For a closed circuit Γ , $A = B$ and we have

$$\varepsilon = \oint_{\Gamma} (\mathbf{E}_{\text{emf}} \cdot d\mathbf{l}) = -\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dt} = RI$$

where R is the total resistance of the circuit.

- Note that the work done on a charge element dq by the force associated with the electromotive field per unit time is

$$dP_{\varepsilon} = dq \left(-\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B} \right) \cdot \mathbf{u},$$

with \mathbf{u} the velocity of the charge with respect to the circuit. The total power generated by the electromotive force is then

$$P_\varepsilon = \oint_{\Gamma} \left(-\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B} \right) \cdot \underbrace{\mathbf{u} dq}_{Idl} = \varepsilon I.$$

8.4 Inductance

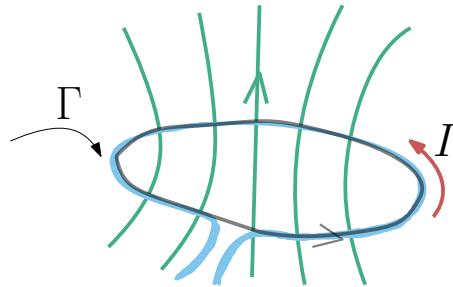
Faraday's law of induction allows us to explain the natural tendency of every conductor to oppose to a change in the current through it. Indeed, a varying current generates a time-varying magnetic field which, in general, will change the magnetic flux through some part of the circuit. An induced electromotive force will then oppose to the change of current.

8.4.1 Self-inductance

A closed circuit carrying a current I generates a magnetic field everywhere in space, and therefore a magnetic flux through the circuit. Consider the closed path Γ of the current and the surface $S(\Gamma)$ enclosed by Γ . We know that the magnetic field \mathbf{B} generated by the circuit is proportional to the current I , so that the magnetic flux is also proportional to I ,

$$\Phi_{S(\Gamma), \mathbf{B}} = \iint_{S(\Gamma)} \mathbf{B}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}) = LI,$$

where L , a positive coefficient which depends only on the geometry of the circuit Γ , is called self-inductance coefficient or, simply, inductance. Its unit is the Henry, with $1 \text{ H} = 1 \text{ Wb A}^{-1}$.



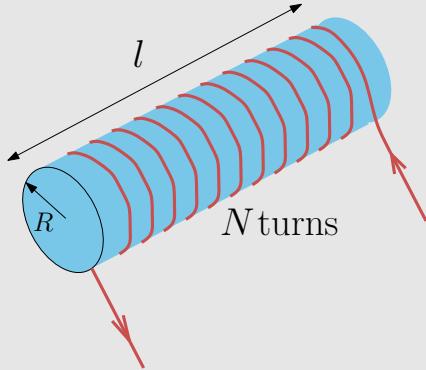
From Faraday's law of induction we know that, if the current through the circuit changes, there will be a self-induced electromotive force ε that will oppose the change in magnetic flux through the circuit, and we have

$$\varepsilon = -\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dt} = -\frac{d\Phi_{S(\Gamma), \mathbf{B}}}{dI} \frac{dI}{dt} = -L \frac{dI}{dt}.$$

The inductance L of a circuit then quantifies how capable it is to oppose a change in the current. The higher L is, the larger the self-induced electromotive force will be.

Example 8.2 - Inductance of a solenoid

What is the inductance of a solenoid of length l , radius $R \ll l$, having N turns? What is its numerical value in Henry for the case $N = 2800$, $R = 5\text{ cm}$ and $l = 0.6\text{ m}$?



Solution

Since $R \ll l$, we may neglect edge effects and use the result for the magnetic field inside an infinitely long solenoid (see chapter 7):

$$\mathbf{B} = B\mathbf{u}_z = \frac{\mu_0 NI}{l}\mathbf{u}_z,$$

where \mathbf{u}_z is the axis of the solenoid. The magnetic flux through the surface enclosed by the circuit of the solenoid is

$$\Phi_B = NB\pi R^2 = \frac{\mu_0 \pi R^2 N^2 I}{l}.$$

Finally, the inductance of this system writes

$$L = \frac{\Phi_B}{I} = \frac{\mu_0 \pi R^2 N^2}{l}.$$

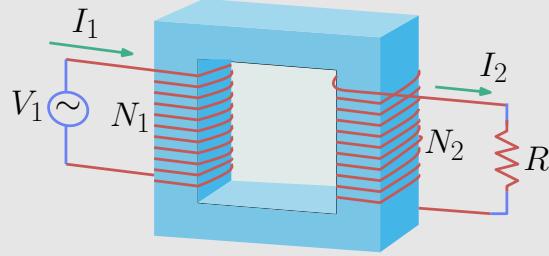
Using the values for the numerical example, we find

$$L = \frac{4\pi^2 10^{-7} \times 0.05^2 \times 2800^2}{0.6} = 0.13\text{ H.}$$

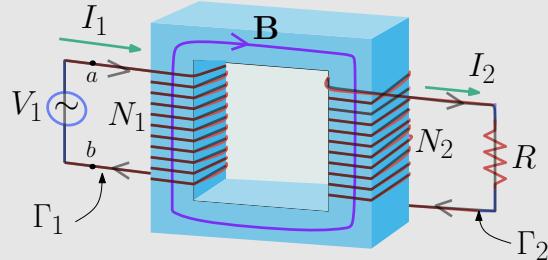
Example 8.3 - Application of induction: The transformer

The transformer is a device used to increase or decrease an alternative electric potential in a circuit. For example, the typical voltage produced in a power plant, of a few kilovolts, is increased to a few hundred kilovolts which is more convenient to transport the electrical energy in transmission lines. Conversely, this voltage is reduced to a few hundred volts for domestic consumption, and may be further reduced (to a few tens of volts) to charge mobile devices, for example.

The transformer is a closed magnetic circuit composed of two coils (typically concentric so as to maximize their magnetic coupling). The first coil, of N_1 turns, is connected at its ends to a sinusoidal potential difference $V_1 = V_0 \cos \omega t$ that generates a sinusoidal current I_1 , which in turns generates a magnetic flux through the second coil, in which an induced current I_2 appears. The coils are winded around an iron core so as to properly confine the magnetic field lines.



Let us define the closed paths Γ_1 and Γ_2 corresponding to the first and secondary coil, respectively. For the first coil, neglecting its internal resistance and suposing no electromotive force inside the generator, generalized Ohm's law (8.9) gives $V(a) - V(b) = V_1(t) = \varepsilon_1$.



If S is the surface enclosed by one turn on the primarily coil, we then have

$$\varepsilon_1 = -N_1 \frac{d\Phi_{S,\mathbf{B}}}{dt} = V_1(t) = V_0 \cos \omega t.$$

The iron core of permeability $\mu \gg \mu_0$ confines the magnetic field lines in between the coils so that in an ideal transformer, the flux through one turn of the secondary coil equals $\Phi_{S,\mathbf{B}}$. Then, the electromotive force ε_2 on Γ_2 is given by

$$\varepsilon_2 = -N_2 \frac{d\Phi_{S,\mathbf{B}}}{dt} = \frac{N_2}{N_1} V_1 \cos \omega t = \frac{N_2}{N_1} \varepsilon_1.$$

The electromotive forces on both coils are thus related by the ratio between the number of turns:

$$\frac{\varepsilon_2}{\varepsilon_1} = \frac{N_2}{N_1}.$$

If this second coil is connected to a load resistor R , then the induced current I_2 is given by

$$I_2(t) = \frac{\varepsilon_2(t)}{R} = \frac{N_2}{N_1} \frac{V_0}{R} \cos \omega t.$$

If the resistance in the coils is neglected, then the total power consumed by the secondary circuit must equal the input power $I_1(t)V_1(t)$, so that

$$I_1(t)V_1(t) = I_2(t)\varepsilon_2(t),$$

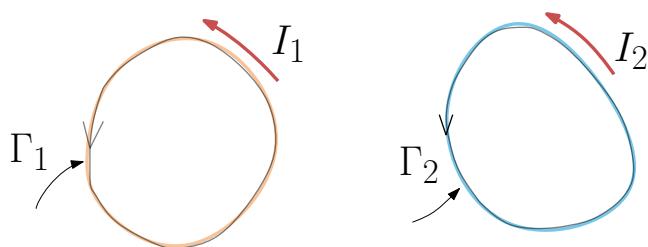
which gives

$$\frac{I_2(t)}{I_1(t)} = \frac{N_1}{N_2} = \frac{\varepsilon_1(t)}{\varepsilon_2(t)}.$$

To transmit electric power from a power plant to a city, a transformer is used to decrease the current (therefore increasing the voltage) so as to reduce losses due to Joule heating, which varies as RI^2 , with I the current in the transmission line, and R its resistance. For example, an increase from 6 to 300 kV will reduce the current at the exit of the transformer by a factor of 50, therefore reducing power loss due to Joule heating by a factor of $50^2 = 2500$.

8.4.2 Magnetic coupling between current loops

Faraday's law of induction allows us to properly describe the magnetic coupling that will generally exist between different parts of a circuit. Indeed, if a varying current flows in a circuit, the magnetic field generated will induce an electromotive force in any loop present nearby. These current loops will, in turn, generate a magnetic field that will influence the original current. Similarly as we did for the capacitance between two conductors, we define the inductance between two circuits as a geometrical coefficient describing the efficiency of their magnetic coupling. Suppose we have two closed paths Γ_1 and Γ_2 corresponding to the arbitrary current loops shown in the figure below.



The flux of the magnetic field \mathbf{B}_1 generated by Γ_1 through any surface $S(\Gamma_2)$ enclosed by the circuit Γ_2 is

$$\Phi_{S(\Gamma_2), \mathbf{B}_1} = \iint_{S(\Gamma_2)} \mathbf{B}_1 \cdot d\mathbf{S}_2 = \iint_{S(\Gamma_2)} (\nabla \times \mathbf{A}_1) \cdot d\mathbf{S}_2,$$

and by the Stokes theorem, this can be rewritten as

$$\Phi_{S(\Gamma_2), \mathbf{B}_1} = \oint_{\Gamma_2} \mathbf{A}_1(\mathbf{x}_2) \cdot d\mathbf{l}_2(\mathbf{x}_2).$$

The vector potential generated by the current I_1 in Γ_1 is given, in Coulomb's gauge, by

$$\mathbf{A}_1(\mathbf{x}_2) = \frac{\mu_0 I_1}{4\pi} \oint_{\Gamma_1} \frac{d\mathbf{l}_1(\mathbf{x}_1)}{|\mathbf{x}_1 - \mathbf{x}_2|}.$$

Finally

$$\Phi_{S(\Gamma_2), \mathbf{B}_1} = I_1 \underbrace{\frac{\mu_0}{4\pi} \oint_{\Gamma_1} \oint_{\Gamma_2} \frac{d\mathbf{l}_1(\mathbf{x}_1) \cdot d\mathbf{l}_2(\mathbf{x}_2)}{|\mathbf{x}_1 - \mathbf{x}_2|}}_M = MI_1,$$

where M exclusively depends on the geometry of the two circuits.

Definition 8.2: Coefficient of mutual induction

The coefficient of mutual induction between two linear circuits Γ_1 and Γ_2 is given by Neumann's formula derived above

$$M = \frac{\mu_0}{4\pi} \oint_{\Gamma_2} \oint_{\Gamma_1} \frac{d\mathbf{l}_1(\mathbf{x}_1) \cdot d\mathbf{l}_2(\mathbf{x}_2)}{|\mathbf{x}_2 - \mathbf{x}_1|} \quad (8.10)$$

and its unit, as for the inductance, is the Henry (H), with $1 \text{ H} = 1 \text{ Wb A}^{-1}$. The magnetic flux generated by Γ_1 through Γ_2 then writes

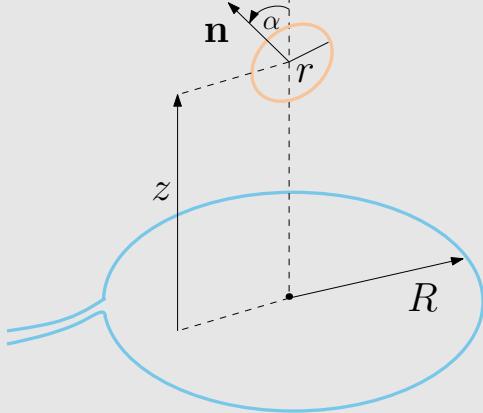
$$\Phi_{S(\Gamma_2), \mathbf{B}_1} = MI_1$$

and since the role of Γ_1 and Γ_2 can be exchanged in the definition of M without changing its value, we conclude that the magnetic flux through Γ_1 generated by Γ_2 is

$$\Phi_{S(\Gamma_1), \mathbf{B}_2} = MI_2.$$

Example 8.4 - Mutual inductance between non-coplanar loops

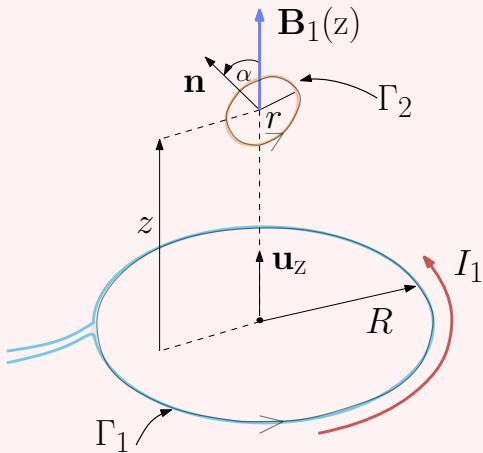
What is the coefficient of mutual induction between the two loops shown in the figure below, of radius r and R , respectively, with $r \ll R$? Both centers lie on the same axis at a distance z , but the normal vectors to the planes containing the loops form an angle α between them.



Solution

A way to obtain the mutual inductance is to suppose that a current flows in the loop Γ_2 of radius r and to calculate the resulting magnetic flux through the surface $S(\Gamma_1)$ enclosed by the loop Γ_1 of radius R . However, this would be too difficult since the magnetic field generated by Γ_2 is highly inhomogeneous over the surface $S(\Gamma_1)$.

Instead, we can suppose that a current I_1 flows in Γ_1 and calculate the magnetic flux over $S(\Gamma_2)$ due to the magnetic field \mathbf{B}_1 generated by I_1 . This flux is easy to estimate since \mathbf{B}_1 is approximately constant over $S(\Gamma_2)$ if $r \ll R$.



The magnetic field on the axis of a circular loop of current was obtained in the last chapter, and applied to Γ_1 , we find

$$\mathbf{B}_1(z) = \frac{\mu_0 I_1 R^2}{2(z^2 + R^2)^{3/2}} \mathbf{u}_z.$$

The magnetic flux through $S(\Gamma_2)$ is then

$$\begin{aligned}\Phi_{S(\Gamma_2), \mathbf{B}_1} &= \iint_{S(\Gamma_2)} \mathbf{B}_1 \cdot \mathbf{n} dS \approx \iint_{S(\Gamma_2)} \left(\frac{\mu_0 I R^2}{2(z^2 + R^2)^{3/2}} \right) \underbrace{\mathbf{u}_z \cdot \mathbf{n}}_{\cos \alpha} dS \\ &= \frac{\mu_0 I_1 R^2 \cos \alpha}{2(z^2 + R^2)^{3/2}} \underbrace{\iint_{S(\Gamma_2)} dS(\mathbf{x})}_{\pi r^2}.\end{aligned}$$

Finally

$$M = \frac{\Phi_{S(\Gamma_2), \mathbf{B}_1}}{I_1} \approx \frac{\mu_0 \pi r^2 R^2 \cos \alpha}{2(z^2 + R^2)^{3/2}}.$$

8.4.3 Energy stored in an inductance

Given the fact that a conductor, through its inductance L , opposes to a change in the magnetic flux, a work must be provided in order to establish a current in the circuit. The power provided must compensate the power of the electromotive force:

$$P = -I\varepsilon = IL \frac{dI}{dt}.$$

If the current increases, $dI/dt > 0$, so that a positive work must be provided by an external source ($P > 0$), and the internal energy U_B stored in the circuit then increases. On the other hand, if the current decreases, $dI/dt < 0$ and the external source takes energy from the circuit, thus reducing its internal energy. The total work provided by the external source to increase the current from 0 to I between $t = 0$ and $t = T$ is

$$W_{\text{ext}} = \int_0^T P(t) dt = \int_0^I LI dI = \frac{1}{2} LI^2,$$

which corresponds to the magnetic energy stored in the circuit

$$U_B = \frac{1}{2} LI^2. \tag{8.11}$$

in analogy to the expression for the electric energy stored in a capacitor

$$U_E = \frac{1}{2} \frac{Q^2}{C}.$$

Example 8.5 - Magnetic energy density in a solenoid

In Ex. 8.2 the inductance L of a solenoid of radius R having N turns and length l was determined:

$$L = \mu_0 \pi R^2 \frac{N^2}{l}.$$

If a current I flows in the solenoid, then the magnetic energy stored in it is given by

$$U = \frac{1}{2} L I^2 = \frac{\mu_0 \pi R^2 N^2 I^2}{2l}.$$

On the other hand, the magnetic field inside an infinitely long solenoid is uniform with magnitude $B = \mu_0 N I / l$, so that in this case the magnetic energy can be written in terms of the magnetic field inside the solenoid:

$$U = \frac{\mu_0 \pi R^2 N^2}{2l} \frac{B^2 l^2}{\mu_0^2 N^2} = l \pi R^2 \frac{B^2}{2\mu_0}.$$

Finally, since $V = \pi R^2 l$ corresponds to the volume of the solenoid, the magnetic energy density stored in the solenoid reads

$$u_B = \frac{U}{V} = \frac{1}{2\mu_0} B^2.$$

8.5 Magnetostatic energy

8.5.1 Energy function for a filamentary circuit in an external magnetic field

In electrostatics, we defined the energy function for a charge q at position $M(\mathbf{r})$ and interacting with a potential $V(\mathbf{r})$. By definition, it is equal to the amount of work

$$W = qV(\mathbf{r})$$

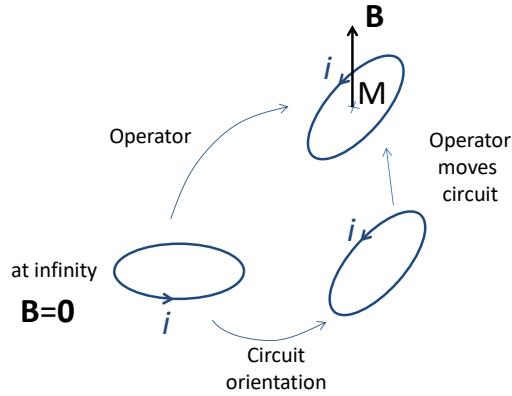
performed by an external operator to bring the charge q from infinity to its final position at M , at potential $V(\mathbf{r})$, while maintaining equilibrium during this process. This means that the external operator is working slowly enough, i.e., quasistatically and in a reversible way, so as not to spill energy in the process. From this result, we inferred that the electrostatic energy for a localized charge distribution of volume density $\rho(\mathbf{r})$ in a volume (Ω) is

$$W = \frac{1}{2} \iiint_{\Omega} \rho(\mathbf{r}) V(\mathbf{r}) d^3 r,$$

and W was shown to be also equal to

$$W = \frac{\epsilon_0}{2} \iiint_{\mathbb{R}^3} \mathbf{E}^2(\mathbf{r}) d^3r.$$

In analogy, we can define an energy function for a filamentary circuit carrying a constant current i , in an external magnetic field $\mathbf{B}(M)$. It is also equal to the amount of work performed by an external operator, working quasistatically and in a reversible way, to give its final orientation to the circuit³ and bring it from infinity to its final position M , where the external magnetic field $\mathbf{B}(M)$ will produce a magnetic flux $\Phi(M)$ through a surface bounded by the circuit.



We will show that the energy function, i.e., the external work of the operator is

$$W = -i\Phi(M), \text{ where } \Phi(M) = \iint_S \mathbf{B} \cdot \mathbf{n} dS$$

We have shown in chapter 7 that the expression for the total force \mathbf{f} on a closed circuit (or dipole moment) can be expressed component-wise by summing the differential elements of Laplace force exerted on the circuit. For the x -component (the expressions would be similar for the y and z -components), the force reads

$$f_x = \int_{\Gamma} id\mathbf{l} \times \mathbf{B} \cdot \mathbf{u}_x$$

Performing a circular permutation on the triple product and using Stokes's theorem, we obtain

$$f_x = \int_{\Gamma} i\mathbf{B} \times \mathbf{u}_x \cdot d\mathbf{l} = \iint_{S(\Gamma)} i\nabla \times (\mathbf{B} \times \mathbf{u}_x) \cdot \mathbf{n} dS,$$

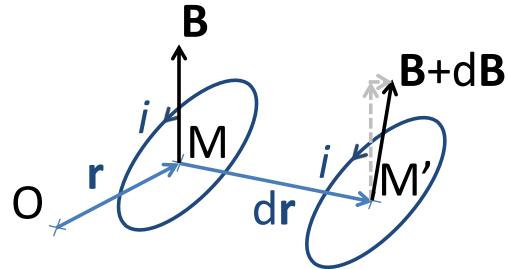
where $S(\Gamma)$ is any surface bounded by Γ . Since $\nabla \times (\mathbf{B} \times \mathbf{u}_x) = \frac{\partial \mathbf{B}}{\partial x}$, we find the force exerted on a circuit in the magnetic field \mathbf{B} :

$$f_x = \iint_{S(\Gamma)} i \frac{\partial \mathbf{B}}{\partial x} \cdot \mathbf{n} dS.$$

³This can always be done first, at infinity, where there is no magnetic field and therefore at no cost in energy since the operator does not have to act against any torque applied on the circuit.

An operator must exert an opposite force $-\mathbf{f} = -f_x \mathbf{u}_x - f_y \mathbf{u}_y - f_z \mathbf{u}_z$ to bring the circuit from infinity to its position in a quasistatic way. For a differential element of displacement $d\mathbf{r} = dx \mathbf{u}_x + dy \mathbf{u}_y + dz \mathbf{u}_z$ of the circuit, the work increment of the operator is $\delta W = -\mathbf{f} \cdot d\mathbf{r}$, which reads

$$\delta W = - \iint_{S(\Gamma)} i \underbrace{\left(\frac{\partial \mathbf{B}}{\partial x} dx + \frac{\partial \mathbf{B}}{\partial y} dy + \frac{\partial \mathbf{B}}{\partial z} dz \right)}_{d\mathbf{B}} \cdot \mathbf{n} dS = - \iint_S i d\mathbf{B} \cdot \mathbf{n} dS = -i \Phi_{S(\Gamma), d\mathbf{B}}.$$



The vector $d\mathbf{B}(M)$ represents the change of magnetic field at M for the differential element of displacement $d\mathbf{M}$. The flux through the circuit is a linear function of the field, hence the flux $\Phi_{S,d\mathbf{B}}$ of $d\mathbf{B}(M)$ through S is nothing but the flux increment $d\Phi$.

$$\delta W = -i \Phi_{S,d\mathbf{B}} = -id\Phi.$$

This is precisely the result we wanted to show: for a reversible transformation, the work increment to move a circuit over $d\mathbf{M}$ appears to be an exact differential, thus

$$W(M) = -i \Phi(M).$$

8.5.2 Forces and torque on a loop in a magnetic field

Consider the closed circuit Γ carrying a steady current i . Laplace's force on a line element $d\mathbf{l}$ is expressed via as the infinitesimal force $d\mathbf{f} = id\mathbf{l} \times \mathbf{B}$. For the entire circuit, the action of the Laplace forces is characterized by the total Laplace force \mathbf{f} and the torque $\boldsymbol{\tau}$. An operator maintaining the circuit at its position and its orientation with respect to the magnetic field must exert an opposite force $-\mathbf{f}$ and torque $-\boldsymbol{\tau}$ on it. In this process, the operator performs a work increment $dW = -id\Phi$ for a differential element of displacement $d\mathbf{M}$ and of rotation characterized by three angle elements $d\alpha, d\beta, d\gamma$ around the axes of the Cartesian basis $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$. The work of Laplace's force (opposite to that of the operator) is then

$$\begin{aligned} -dW &= id\Phi = \mathbf{f} \cdot d\mathbf{M} + \boldsymbol{\tau} \cdot d\boldsymbol{\omega} \\ &= f_x dx + f_y dy + f_z dz + \tau_\alpha d\alpha + \tau_\beta d\beta + \tau_\gamma d\gamma. \end{aligned}$$

The work function was already shown to be $W(M) = -i\Phi(M)$. The flux $\Phi(M)$ depends on the position of the center of mass of the circuit and of the orientation of the circuit, given by angular coordinates α, β, γ .

$$\Phi(\underbrace{x, y, z}_{\text{position}}, \underbrace{\alpha, \beta, \gamma}_{\text{orientation}})$$

The force and torque on the circuit are thus retrieved component-wise from the energy function:

$$f_x = -\frac{\partial W}{\partial x} = i \frac{\partial \Phi}{\partial x},$$

$$\tau_\alpha = -\frac{\partial W}{\partial \alpha} = i \frac{\partial \Phi}{\partial \alpha},$$

and relations for the other components are similar.

8.5.3 Magnetic free energy

In this section, we derive the magnetic energy for a set of conductors carrying currents. The conductors are in interaction since each generates a magnetic field, which induces a magnetic flux in all other conductors. We then interpret this magnetic energy as the free energy of the set of conductors. The terminology of *free energy* simply reflects the fact that the magnetic system of conductors is treated as a thermodynamic system exchanging electric work, and possibly heat, with its environment. The free energy then serves as a potential allowing us to derive equilibrium properties of the system and forces acting on it. The concept of free energy is reviewed in section 8.6 together with reminders of thermodynamics.

- **CONTINUOUS DISTRIBUTION OF CURRENTS**

Consider a set of metallic conductors $(\Gamma_l), l = 1, \dots, n$, e.g. filamentary circuits (loops) carrying currents i_l , occupying a finite volume of space. We treat them as a continuous distribution of currents and will adapt the findings to the case of filamentary circuits below. At any point \mathbf{r} of space, the current density vector $\mathbf{J}(\mathbf{r})$ and the vector potential $\mathbf{A}(\mathbf{r})$ are supposed to be initially zero and we are looking for the energy required to establish a steady state regime characterized by a permanent current density $\mathbf{J}(\mathbf{r})$ and vector potential $\mathbf{A}(\mathbf{r})$. We know (chapter 7) that these quantities must satisfy Poisson's equation

$$\Delta \mathbf{A}(\mathbf{r}) = -\mu_0 \mathbf{J}(\mathbf{r}).$$

In order to establish the final regime, we assume that a set of generators, electrically connected to the conductors but external to the system, can set charge carriers in motion to establish the current density within the conductors. The road to the final state is slow enough so as to maintain the steady state equilibrium all along the way.

Linearity of the Poisson equation means that any intermediate steady state regime between the initial $(\mathbf{0}, \mathbf{0})$ and final state $(\mathbf{J}(\mathbf{r}), \mathbf{A}(\mathbf{r}))$ is characterized by a current $\lambda \mathbf{J}(\mathbf{r})$ and vector potential $\lambda \mathbf{A}(\mathbf{r})$, where λ is a coefficient such that $0 \leq \lambda \leq 1$.

At time t , the vector potential is therefore, at any point M , $\lambda(t)\mathbf{A}(\mathbf{r})$, whose time dependency, even if it is slow, induces an electromotive field $\mathbf{E}_{\text{emf}} = -\frac{d\lambda}{dt}\mathbf{A}(\mathbf{r})$. The generators must then act against this field and apply an opposite electric field on the charge carriers of the conductors in order to maintain the already established steady state regime characterized by $\lambda\mathbf{J}(\mathbf{r})$. The generators then give a power p per unit volume to the conductors

$$p = -\mathbf{E}_{\text{emf}} \cdot (\lambda\mathbf{J}) = \lambda \frac{d\lambda}{dt} \mathbf{A}(M) \cdot \mathbf{J}(\mathbf{r}).$$

that is the amount of work per unit volume, dw , done by the external generators for the system of conductors to change by an amount of $d\lambda$ is

$$dw = pdt = \mathbf{A}(\mathbf{r}) \cdot \mathbf{J}(\mathbf{r}) \lambda d\lambda.$$

Summing from $\lambda = 0$ to $\lambda = 1$, the total amount of work density provided to the conductors is

$$w(\mathbf{r}) = \frac{1}{2} \mathbf{A}(\mathbf{r}) \cdot \mathbf{J}(\mathbf{r}).$$

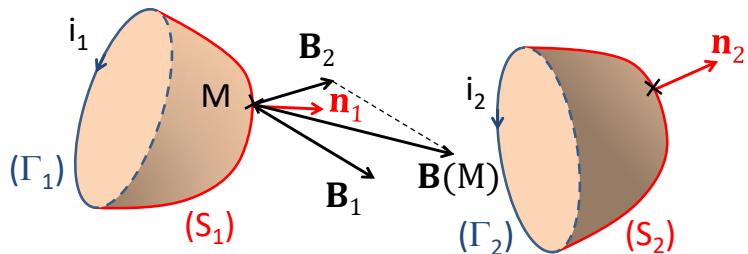
In thermodynamics, we know that the work provided by external operators during a quasistatic reversible transformation is equal to the change⁴ in the free energy of the system

$$W = \Delta F = F_{\text{final}} - F_{\text{initial}},$$

hence $w(\mathbf{r})$ corresponds to the free energy density of the system (assuming $F_{\text{initial}} = 0$). Now for the entire volume (Ω) occupied by the conductors, the free energy of the system reads

$$F = \frac{1}{2} \iiint_{\Omega} \mathbf{A}(\mathbf{r}) \cdot \mathbf{J}(\mathbf{r}) d^3 r. \quad (8.12)$$

- CASE OF FILAMENTARY CIRCUITS



⁴ Δ is here a difference, not to be confused with the Laplacian operator.

Equation (8.12) is easy to apply to the case of n filamentary circuits (Γ_l), $l = 1, \dots, n$, carrying currents i_l , as shown in the figure for the case $n = 2$. The volume integral becomes a line integral after changing the current $\mathbf{J}(\mathbf{r})d^3r$ into $id\mathbf{l}$. The free energy then reads

$$F = \frac{1}{2} \sum_{l=1}^n \oint_{\Gamma_l} \mathbf{A}(M) \cdot i_l d\mathbf{l},$$

where $d\mathbf{l}_l$ denotes the infinitesimal line element vector that is tangent at each point of the circuit l . We can then apply Ostrogradsky's theorem to the circulation integrals, leading to

$$F = \frac{1}{2} \sum_{l=1}^n i_l \underbrace{\iint_{S_l} \mathbf{B}(\mathbf{r}) \cdot \mathbf{n}_l dS_l}_{\Phi_l}$$

and finally,

$$F = \frac{1}{2} \sum_{l=1}^n i_l \Phi_l$$

where Φ_l denotes the flux of the magnetic field through Γ_l .

Since the magnetic field \mathbf{B} is the superposition of the fields generated by the set of circuit, the magnetic flux through Γ_l is contributed by the field \mathbf{B}_l generated by circuit l itself, and by the fields \mathbf{B}_k , with $k \neq l$ generated by all other circuits. Therefore, the flux is expressed as

$$\Phi_l = \underbrace{L_l i_l}_{\text{self-flux}} + \underbrace{\sum_{k \neq l} M_{lk} i_k}_{\text{induced flux}},$$

where we recognize the self-flux and the fluxes induced by other circuits. The coefficients L_l and M_{lk} for $k \neq l$, are the self inductance and mutual inductance coefficients, respectively. This equation can be rewritten in matrix form as

$$\begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_n \end{bmatrix} = \underbrace{\begin{bmatrix} L_1 & \dots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{n1} & \dots & L_n \end{bmatrix}}_{\text{inductance matrix } [\mathcal{L}]} \begin{bmatrix} i_1 \\ \vdots \\ i_n \end{bmatrix}$$

where self-inductances appear on the diagonal and mutual inductances in other places of the inductance matrix matrix $[\mathcal{L}]$. The inductance matrix exhibits specific properties reviewed below.

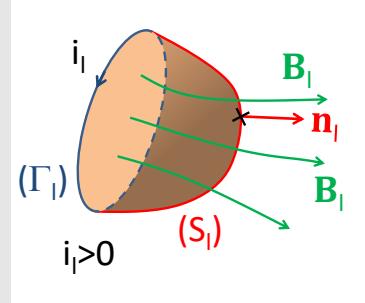
Example 8.6 - Properties of the inductance matrix

- The inductance matrix is symmetric: $M_{lk} = M_{kl}$. This directly follows from Neuman's formula, already demonstrated (See 8.10). For a system of two circuits, we

have $M_{12} = M_{21} = M$ and

$$\begin{bmatrix} \Phi_2 \\ \Phi_2 \end{bmatrix} = \begin{bmatrix} L_1 & M \\ M & L_2 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix}$$

- Diagonal coefficients (self-inductances) are positive: $L_l > 0$



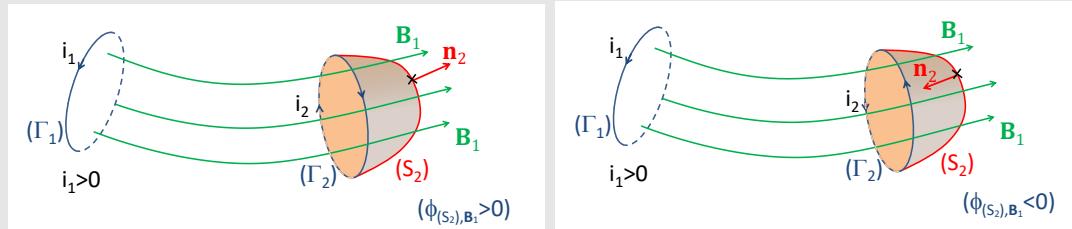
This follows directly from the orientation of a circuit. The direction of the current i_l defines a north and a south pole for the circuit Γ_l . The magnetic field \mathbf{B}_l generated by the circuit points from the south pole to the north pole, which is exactly the orientation of the normal \mathbf{n}_l to the surface $S_l(\Gamma_l)$ bounded by the circuit. Therefore the self-induced flux $\Phi_{S_l(\Gamma_l), \mathbf{B}_l}$ has the same sign as the current and from $\Phi_{S_l(\Gamma_l), \mathbf{B}_l} = L_l i_l$, we deduce that $L_l > 0$, for all l .

- The inductance matrix $[\mathcal{L}]$ is a positive definite matrix (see remark below). This means that $\det[\mathcal{L}] > 0$. For a system of two conductors, this yields an upper bound to the mutual inductance coefficient: $M^2 < L_1 L_2$.
- The sign of mutual inductances depends on the circuit orientation. The figures below show the same system of two circuits with two different orientations of Γ_2 . In the first case, the flux $\Phi_{S_2(\Gamma_2), \mathbf{B}_1}$ (of \mathbf{B}_1 through $S_2(\Gamma_2)$) is positive and so is the mutual inductance coefficient.

$$\Phi_{S_2(\Gamma_2), \mathbf{B}_1} > 0, \quad \Phi_{S_2(\Gamma_2), \mathbf{B}_1} = M i_1 \quad \Rightarrow \quad M > 0.$$

In the second case, the opposite orientation of Γ_2 leads to the opposite sign for $\Phi_{S_2(\Gamma_2), \mathbf{B}_1}$ and for the mutual inductance coefficient.

$$\Phi_{S_2(\Gamma_2), \mathbf{B}_1} < 0, \quad \Phi_{S_2(\Gamma_2), \mathbf{B}_1} = M i_1 \quad \Rightarrow \quad M < 0.$$



Remark:

- The self-field \mathbf{B}_l is not defined for a filamentary circuit. Biot-Savart's law is indeed a divergent integral when it is applied to look for the magnetic field on a point of the filamentary distribution itself. It is therefore not possible to calculate the self-inductance coefficient for a filamentary circuit. This problem does not arise with surface or volume current distributions, therefore this singularity can be avoided by using a small but finite wire radius to calculate the self-inductance of a filamentary circuit. Mutual inductance coefficients can always be calculated for filamentary circuits without problem.
- The positive-definiteness of the inductance matrix will be clear in the next section where the magnetic free energy is shown to be expressed as $F = \frac{1}{2\mu_0} \iiint_{\mathbb{R}^3} \mathbf{B}^2 d^3r$. This quantity is always positive. The magnetic free energy was expressed earlier as a quadratic form of the current vector (i_1, \dots, i_n) : $F = \frac{1}{2} \sum_{i=1}^n i_l [L_{ll} i_l + \sum_{k \neq l} M_{lk} i_k]$, whose coefficients are those of the inductance matrix. Thus the inductance matrix is positive-definite as the corresponding quadratic form.

8.5.4 Another expression for the volume density of magnetic free energy

From the free energy (8.12), we can first extend the integration volume to the entire space \mathbb{R}^3 since $\mathbf{J} = 0$ outside (Ω) , and then replace the current density using Maxwell-Ampère's equation $\mathbf{J} = \nabla \times \frac{\mathbf{B}}{\mu_0}$. We obtain

$$F = \iiint_{\mathbb{R}^3} \frac{1}{2\mu_0} \mathbf{A} \cdot \nabla \times \mathbf{B} d^3r$$

Now we use a general relation of vector calculus for the divergence of the cross product between two vectors \mathbf{A} and \mathbf{B} :

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = -\mathbf{A} \cdot \nabla \times \mathbf{B} + \mathbf{B} \cdot \nabla \times \mathbf{A}.$$

Applying it to the vector potential and the magnetic field, and using $\mathbf{B} = \nabla \times \mathbf{A}$, we find

$$\mathbf{A} \cdot \nabla \times \mathbf{B} = -\nabla \cdot (\mathbf{A} \times \mathbf{B}) + \mathbf{B}^2,$$

and

$$F = -\frac{1}{2\mu_0} \iiint_{\mathbb{R}^3} \nabla \cdot (\mathbf{A} \times \mathbf{B}) d^3r + \frac{1}{2\mu_0} \iiint_{\mathbb{R}^3} \mathbf{B}^2 d^3r,$$

whose first term on the right-hand side can be transformed by Ostrogradsky's theorem as

$$F = -\frac{1}{2\mu_0} \iint_{S \rightarrow \infty} (\mathbf{A} \times \mathbf{B}) d^2r + \frac{1}{2\mu_0} \iiint_{\mathbb{R}^3} \mathbf{B}^2 d^3r.$$

The surface integral does not contribute since $\mathbf{A} \sim r^{-1}$ and $\mathbf{B} \sim r^{-2}$ at infinity. Finally

$$F = \frac{1}{2\mu_0} \iiint_{(\mathbb{R}^3)} \mathbf{B}^2 d^3r,$$

from which we identify the volume density of magnetic energy

$$f = \frac{B^2}{2\mu_0}.$$

8.5.5 Free energy vs free enthalpy

Going back to the derivation of Eq. 8.12, we note that as the change from the permanent regime (\mathbf{J}, \mathbf{A}) to the neighbor state $(\mathbf{J} + d\mathbf{J}, \mathbf{A} + d\mathbf{A})$ involves the generation of an electric field $d\mathbf{A}/dt$ by the generators⁵, the increment of free energy density during this transformation is

$$df = \mathbf{J}(M) \cdot d\mathbf{A}(M), \quad (8.13)$$

which shows that the natural variable for the volume density of free energy $f(\mathbf{A})$ is \mathbf{A} , the vector potential.⁶ Poisson's equation for the vector potential implies a linear relation between \mathbf{J} and \mathbf{A} (say $\mathbf{A} = \lambda\mathbf{J}$), which serves as an equation of state. The differential for the free energy density can be integrated by using a single variable, say \mathbf{J} . From the equation of state, we can write $d\mathbf{A} = \lambda d\mathbf{J}$ and we retrieve the result

$$f = \int \lambda \mathbf{J} \cdot d\mathbf{J} = \lambda \frac{\mathbf{J}^2}{2} = \frac{1}{2} \mathbf{J} \cdot \mathbf{A}.$$

Note that the same result can be found if the differential df has been integrated with respect to its natural variable \mathbf{A} rather than \mathbf{J} , but the key point is that these variables are linked by an equation of state.⁷

For filamentary circuits, the free energy was expressed as

$$F = \frac{1}{2} \sum_l i_l \Phi_l,$$

where the set of currents i_l correspond to the variable \mathbf{J} and the set of fluxes Φ_l correspond to the variable \mathbf{A} . Currents and fluxes are conjugated variables, in the same way as \mathbf{J} and \mathbf{A} , i.e., they are linked by an equation of state, that we have shown to be

$$\Phi_l = \sum_{k=1}^n [\mathcal{L}]_{lk} i_k.$$

This implies that while the value of F is expressed by means of products $i_l \Phi_l$ of conjugated variables i_l and Φ_l , the free energy of a set of filamentary conductors can be expressed as a function of a single variable of each pair (i_l, Φ_l) :

$$F = \frac{1}{2} \sum_{l=1}^n \sum_{k=1}^n i_l [\mathcal{L}]_{lk} i_k.$$

⁵this electric field is opposite to the electromotive field.

⁶In the same way in thermodynamics, for a gas undergoing a quasistatic transformation, the volume v and entropy S are the natural variables for the energy $U(v, S)$ as shown by the expression of the exact differential $dU = -pdv + TdS$. The free energy $F(v, T)$ is a function of volume and temperature T with $dF = -pdv - SdT$.

⁷In the same way, pressure and volume are conjugated variables in thermodynamics, linked by the equation of state $pv \propto T$ for an ideal gas. In a transformation between two states, the work is calculated after expressing one of the conjugated variable as a function of the other via the equation of state.

Since the flux Φ_l in each pair (i_l, Φ_l) constitutes the natural variable for the free energy, as reflected by the differential for F

$$dF = \sum_l i_l d\Phi_l,$$

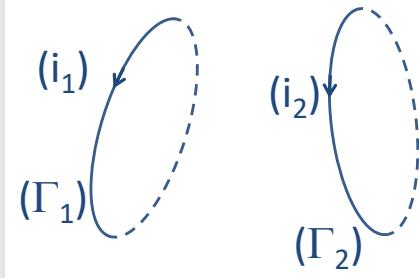
which is analogous to Eq. (8.13), it could be useful to express F as a function of Φ_l rather than i_l , which is easily done by means of the inverse of the inductance matrix⁸

$$F(\Phi_1, \dots, \Phi_n) = \frac{1}{2} \sum_{l=1}^n \sum_{k=1}^n \Phi_l [\mathcal{L}]_{lk}^{-1} \Phi_k.$$

This expression is useful to derive the force and torque applied to an element of the set of conductors from the energy as illustrated below.

Example 8.7 - Free energy for a set of two filamentary circuits

Consider two oriented circuits connected to generators which establishes the currents in each circuit. What is the energy to establish a set of currents in two filamentary circuits?



Any change of current will induce a change in the magnetic flux through each circuit. An electromotive force $-d\Phi_l/dt$, $l = 1, 2$, will thus appear in the circuits. The generators must act against this emf by producing an opposite emf $e_{lg} = +d\Phi_l/dt$.

During dt , generators provide the (free) energy:

$$dF = \underbrace{\frac{d\Phi_1}{dt}}_{e_{1g}} i_1 dt + \frac{d\Phi_2}{dt} i_2 dt = i_1 d\Phi_1 + i_2 d\Phi_2.$$

The fluxes can be expressed as functions of currents from the matrix relation $[\Phi_1, \Phi_2]^\dagger = [\mathcal{L}] [i_1, i_2]^\dagger$, where \dagger denotes transposition, leading to

$$\begin{aligned} dF &= i_1 (L_1 di_1 + M di_2) + i_2 (M di_1 + L_2 di_2) \\ &= L_1 i_1 di_1 + M(i_1 di_2 + i_2 di_1) + L_2 i_2 di_2. \end{aligned}$$

The free energy F is finally obtained by summing from $(0, 0)$ to the final state of currents (i_1, i_2) . The final value for F reads

$$F = \frac{1}{2} L_1 i_1^2 + M i_1 i_2 + \frac{1}{2} L_2 i_2^2.$$

⁸It exist as $[\mathcal{L}]$ is a positive definite matrix.

Since the natural variables for the free energy of a set of filamentary conductors are the magnetic fluxes and not the currents, it is tricky to derive forces and torque applied to a specific circuit from the expression of F as a function of currents. Indeed, one would like to use a similar formulation as for the energy function in section 8.6.2, that is, calculate the torque components, for example, from $\tau_\alpha = -\partial F/\partial \alpha$. This is possible provided the derivation is performed at constant fluxes $\tau_\alpha = -\partial F/\partial \alpha|_{\Phi_l}$, but the generators keep the currents i_l fixed in the circuits, not the fluxes. Another potential is therefore required to facilitate this task, and to ensure that derived forces and torques are compatible with states of equilibrium involving the fixed currents maintained by the generators. This potential is obtained by a Legendre transform of the free energy, which results in the free enthalpy whose volume density g is

$$g = f - \mathbf{A} \cdot \mathbf{J}.$$

Differentiating this definition leads to

$$dg = \underbrace{\frac{df}{\mathbf{J} \cdot d\mathbf{A}}}_{-\mathbf{A} \cdot d\mathbf{J}} - \mathbf{A} \cdot d\mathbf{J} - \mathbf{J} \cdot d\mathbf{A} \Rightarrow dg = -\mathbf{A} \cdot d\mathbf{J},$$

which shows that the natural variable for the free enthalpy is indeed the current density \mathbf{J} . Similarly, for a set of filamentary conductors, the free enthalpy is obtained from a Legendre transform of the free energy as

$$G = F - \sum_l i_l \Phi_l.$$

Differentiating, we find

$$dG = \underbrace{\frac{dF}{\sum_l i_l d\Phi_l}}_{-\sum_l (i_l d\Phi_l + \Phi_l di_l)} - \sum_l (i_l d\Phi_l + \Phi_l di_l) \Rightarrow dG = -\sum_l \Phi_l di_l,$$

which shows that currents i_l are the natural variables of the free enthalpy.

Example 8.8 - Free enthalpy for two filamentary circuits

We can express the free enthalpy as the Legendre transform of the free energy. For a set of two conductors, this yields $G(i_1, i_2) = F - i_1 \Phi_1 - i_2 \Phi_2$. From the differential of the free energy $dF = i_1 d\Phi_1 + i_2 d\Phi_2$, we find that of the free enthalpy

$$dG = -\Phi_1 di_1 - \Phi_2 di_2.$$

Incidentally, the fact that the free enthalpy is a state function provides an alternative to Neuman's formula showing that the inductance matrix is symmetric. The differential dG is indeed exact and therefore satisfies the Schwarz property: $\frac{\partial \Phi_1}{\partial i_2} = \frac{\partial \Phi_2}{\partial i_1}$, which is simply equivalent to the symmetry of mutual inductance coefficients $M_{12} = M_{21} = M$.

We finally obtain the expression for the free enthalpy by integration with respect to the currents,

after using the equation of state $[\Phi_1, \Phi_2]^\dagger = [\mathcal{L}][i_1, i_2]^\dagger$ in the expression of dG :

$$G(i_1, i_2) = -\frac{1}{2}L_1i_1^2 - Mi_1i_2 - \frac{1}{2}L_2i_2^2.$$

If angle α_1 denotes the rotation angle of circuit 1 with respect to an axis, the torque applied to circuit 1 with respect to this axis is expressed as

$$\tau_{\alpha_1} = -\left.\frac{\partial G}{\partial \alpha_1}\right|_{i_1, i_2} = \frac{1}{2}\frac{\partial L_1}{\partial \alpha_1}i_1^2 + \frac{\partial M}{\partial \alpha_1}i_1i_2 + \frac{1}{2}\frac{\partial L_2}{\partial \alpha_1}i_2^2.$$

The currents maintained by the generators are held constant.

To obtain a torque consistent with this result from the free energy F , one must first express the free energy as a function of its natural variables, Φ_1 and Φ_2 , derive the torque from $\tau_{\alpha_1} = -\left.\frac{\partial F}{\partial \alpha_1}\right|_{\Phi_1, \Phi_2}$ and finally use the equation of state to express the result as a function of i_1 and i_2 . Naturally both methods give the same result but this illustrates the care that must be taken on the natural variables of potential functions and the usefulness of the free enthalpy.

8.6 Reminders on energy and potentials in thermodynamics

The goal of this section is to serve as reminders on the concepts of thermodynamics that are used in the electrodynamics course. In particular we remind the definition and physical interpretation of the (Helmholtz) free energy and free enthalpy (also called Gibbs free energy).

- **FIRST PRINCIPLE**

In a mechanical conservative system, the total energy is shared between kinetic and potential energy. Equilibrium states are generally found by looking for minima of the potential energy. For thermodynamical systems, the equivalent principle of energy conservation is the first principle, which states that in a transformation from initial state i to final state f , the variation of energy of the system is equal to the energy received in the form of work W performed by external forces acting on the system and in the form of heat Q exchanged with the environment

$$\Delta U = U(f) - U(i) = W + Q.$$

The energy function U is a state function, that depends only on the state of the system, and therefore, its variation only depends on the initial and final states but not on the path followed between.

• SECOND PRINCIPLE

The second principle states that if the system is in thermal contact with the environment at a constant temperature T , there exist an other state function, the entropy S , such that the exchanged heat during the transformation from i to f satisfies

$$Q = \int_i^f \delta Q \leq T(S(f) - S(i)). \quad (8.14)$$

This means that there is an upper limit to the amount of heat the system can receive from the environment, and this limit determines the entropy of the system.

• THE FREE ENERGY

Introducing inequality (8.14) into the first principle leads to

$$U(f) - U(i) - T(S(f) - S(i)) \leq W,$$

which means that the amount of work that the system can receive from external forces during the transformation is an upper limit for the variation of the function $U - TS$. This function is another state function, the free energy $F = U - TS$, also called the Helmholtz free energy, which satisfies

$$F(f) - F(i) \leq W. \quad (8.15)$$

Now if the transformation is reversible, the equality sign holds in (8.15), that is, in a reversible transformation between two states which have a temperature equal to that of the environment, the work of external forces received by the system is equal to the variation of free energy of the system. In other words, for reversible transformations with heat exchange with the environment at temperature T , the free energy in a thermodynamical system plays a role analogous to that of energy for mechanical systems. In addition, for systems that do not exchange work with the environment ($W = 0$)⁹, the free energy cannot increase and consequently, reaches a minimum when the system is in a state of stable equilibrium, similarly to a dynamically isolated mechanical system that exchange no work from the environment and that has equilibrium states corresponding to minima of its potential energy .

• NATURAL VARIABLES OF STATE FUNCTIONS

The free energy plays the role of a potential in thermodynamics. As any potential, it is used to find minima, and thus equilibrium states of the system. In this aim, it should be expressed as a function of its natural variables, which denote the independent thermodynamical state variables describing the system. Since thermodynamical systems are usually governed by a an equation of state, that is, a relation between the thermodynamical variables, the number of independent variables is reduced accordingly. For example an ideal gas satisfies the law of ideal gases $pv \propto T$ connecting the pressure p to the temperature T and volume v . Hence, a set of two independent variables, e.g., (v, T) or (p, T) , is sufficient to define the state of a system and the dependence of state functions.

Although the choice of independent variables is arbitrary, it is convenient to use a set of independent variables suited to the system transformation under investigation. The set (v, T) is convenient for

⁹This is the case, for example, for a gas undergoing an isochore transformation, that is, a gas of uniform pressure enclosed in a container of fixed volume.

studying isochore and isothermal transformations. The set (p, T) is convenient for studying isobar and isotherm transformations. For a gas, the free energy $F(v, T)$ is appropriately expressed as a function of the volume v and temperature T . For this reason, it is also called the potential for isotherm and isochore transformations. v and T are called the *natural variables* of the free energy. Applied to infinitesimal reversible transformations, the principles of thermodynamics allows us to express differential forms of state functions like the free energy. The state function can then be expressed as a function of its natural variables by integration, using the fact that mathematically, a differential form for a state function corresponds to an exact differential.

• EXACT DIFFERENTIALS

Consider a differential form dz of two independent variables x and y :

$$dz = P(x, y)dx + Q(x, y)dy$$

This differential form is exact (or perfect) if it is the differential of a function of x and y . If that is the case, P and Q must satisfy

$$\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x} \quad (8.16)$$

and vice-versa. Equation (8.16) is the Schwartz property. When it is satisfied, it is possible to integrate the differential form and find a function of x and y that satisfy this equation. In physics, such functions are called state functions. For example, for an ideal gas, the differential form for the free energy reads

$$dF = -pdv - SdT,$$

and the equation of state for the pressure $p(v, T) = nRT/v$ and energy $U(T) = C_vT$ (it is a property of ideal gases that U is independent of v), allows for finding the entropy $S(v, T)$ and in turn the free energy $F(v, T)$

• THE FREE ENTHALPY

The free energy is not the only thermodynamic potential that can be defined. Taking the example of a gas, the free energy $F(v, T)$ is useful to find equilibrium states for isothermal and isochore transformations at constant volume and temperature. If in contrast a transformation is isothermal and isobaric, that is, if it occurs at constant temperature T and pressure p , with a change of volume from $v(i)$ to $v(f)$, the work of external forces is that of pressure forces $-p\Delta v$, leading to the inequality

$$W = -p(v(f) - v(i)) \geq F(f) - F(i)$$

Defining a new state function G , the free enthalpy, or Gibbs free energy,

$$G = F + pv = U - TS + Pv,$$

we see that

$$G(f) \leq G(i)$$

In other words the free enthalpy G cannot increase for isothermal and isobaric transformations. Therefore it plays the role of a thermodynamic potential at constant pressure. For these transformations, minima of the free enthalpy are stable equilibrium states.

Differentiating the definition of the free enthalpy, we find

$$dG = dF + pdv + vdp$$

and introducing $dF = -pdv - SdT$ leads to

$$dG = vdp - SdT$$

which shows that the natural variables for the free enthalpy $G(p, T)$ are the pressure p and temperature T , that is, $G(p, T)$ is the thermodynamic potential for isobar and isothermal transformations.

8.7 Summary and essential formulas

- When a circuit (or a part of it) moves with velocity \mathbf{v} in the presence of a time-varying magnetic field \mathbf{B} , an electromotive field \mathbf{E}_{emf} acts on the electrons of the circuit generating an induced current. This field is given by

$$\mathbf{E}_{\text{emf}} = \mathbf{v} \times \mathbf{B} - \frac{\partial \mathbf{A}}{\partial t}.$$

The term $\mathbf{v} \times \mathbf{B}$ comes from the motion of the conductor in the magnetic field, whereas the term $-\frac{\partial}{\partial t} \mathbf{A}$ is an electric field generated by the time-dependent magnetic field. This second term is at the heart of a profound relationship between time-dependent electric and magnetic fields.

- The circulation of the electromotive field around a closed path Γ is called the electromotive force, and it is related to the change of magnetic flux enclosed by Γ according to Faraday's law

$$\varepsilon = \oint_{\Gamma} \mathbf{E}_{\text{emf}} \cdot d\mathbf{l} = -\frac{d}{dt} \Phi_{S(\Gamma), \mathbf{B}},$$

where

$$\Phi_{S(\Gamma), \mathbf{B}} = \iint_{S(\Gamma)} \mathbf{B} \cdot \mathbf{n} dS$$

is the magnetic flux through any surface whose bounding curve is Γ .

- The total derivative of the magnetic flux $\Phi_{S(\Gamma), \mathbf{B}}$ with respect to time is

$$-\frac{d}{dt} \Phi_{S(\Gamma), \mathbf{B}} = -\left(\mathbf{v} \cdot \nabla + \frac{\partial}{\partial t} \right) \Phi_{S(\Gamma), \mathbf{B}}$$

and the Stokes theorem shows that each term in the electromotive field corresponds to one of the two terms of the time derivative:

$$\oint_{\Gamma} (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = \underbrace{-(\mathbf{v} \cdot \nabla) \Phi_{S(\Gamma), \mathbf{B}}}_{\begin{array}{c} \text{advection term, due to} \\ \text{the motion of the circuit} \end{array}}$$

and

$$\oint_{\Gamma} \left(-\frac{\partial \mathbf{A}}{\partial t} \right) \cdot d\mathbf{l} = \underbrace{-\frac{\partial}{\partial t} \Phi_{S(\Gamma), \mathbf{B}}}_{\begin{array}{c} \text{due to explicit time-dependence} \\ \text{of the magnetic field} \end{array}}$$

- The circulation law for the electric field is modified in the presence of a time-dependent magnetic field. The differential form of Faraday's law writes:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

so that the electric field is no longer conservative in the time-varying regime. Since $\mathbf{B} = \nabla \times \mathbf{A}$, we see that $\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0$ and therefore $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla V$ where V is the usual electric potential in Coulomb's gauge ($\nabla \cdot \mathbf{A} = 0$). Then,

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}.$$

The first term is the conservative term due to the electrostatic potential (charge separation), whereas the second term, which is not conservative, corresponds to the contribution of the induced electric field to the electromotive field.

- Ohm's law is generalized for the case of a non-conservative electric field. The potential difference between points A and B of a conductor reads

$$V_A - V_B = R_{AB}I - \varepsilon_{AB},$$

where I is the current flowing from A to B , R_{AB} is the total resistance of the segment AB and ε_{AB} the electromotive force between A and B .

- The forces and torques acting on a circuit (loop) carrying a fixed current i in a magnetic \mathbf{B} field can be derived from an energy function

$$W = -i\Phi, \quad \text{with} \quad \Phi = \iint_S \mathbf{B} \cdot \mathbf{n} dS,$$

using the following formulas, component-wise:

$$f_x = -\frac{\partial W}{\partial x} = i \frac{\partial \Phi}{\partial x}, \quad \tau_\alpha = -\frac{\partial W}{\partial \alpha} = i \frac{\partial \Phi}{\partial \alpha}.$$

- For a current distribution of density \mathbf{J} in a volume Ω , the density of free energy per unit volume and the free energy read

$$f = \frac{1}{2} \mathbf{J} \cdot \mathbf{A} \quad \text{and} \quad F = \frac{1}{2} \iiint_{\Omega} \mathbf{J} \cdot \mathbf{A} d^3 r,$$

respectively, or

$$f = \frac{1}{2\mu_0} B^2 \quad \text{and} \quad F = \frac{1}{2} \iiint_{(\mathbb{R}^3)} \frac{B^2}{\mu_0} d^3 r.$$

- For a set of filamentary circuits carrying currents i_l , connected with generators that maintain the currents at equilibrium, the equation of state linking fluxes and currents in each circuit reads (example with two circuits):

$$\begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} = \begin{bmatrix} L_1 & M \\ M & L_2 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix},$$

where the diagonal coefficients define the self-inductances and the off diagonal coefficients define the mutual inductance. Self-inductances are positive. The inductance matrix is a symmetric positive definite matrix.

- Mutual inductance coefficients satisfy Neuman's formula (for any set of two circuits):

$$M_{12} = \int_{(\Gamma_1)} \int_{(\Gamma_2)} \frac{\mu_0}{4\pi} \frac{d\mathbf{l}_2 \cdot d\mathbf{l}_1}{r}.$$

- The free energy is a state function that represents the work of external forces to establish a set of currents. It is therefore positive. For a set of two circuits, its differential reads

$$dF = i_1 d\Phi_1 + i_2 d\Phi_2,$$

and its value

$$F = \frac{1}{2} L_1 i_1^2 + M i_1 i_2 + \frac{1}{2} L_2 i_2^2.$$

For a single circuit carrying a current i , the stored magnetic energy is

$$F = \frac{1}{2} L i^2.$$

- The force and torque applied on a circuit (say Γ_k) within a set of circuits carrying currents i_l , $l = 1 \dots n$ can be derived from the free energy $F(\Phi_l, x_l, \alpha_l, \dots)$

$$f_{x, \Gamma_k} = - \left. \frac{\partial F}{\partial x_k} \right|_{\Phi_l, \dots}, \quad \tau_\alpha = - \left. \frac{\partial F}{\partial \alpha_k} \right|_{\Phi_l, \dots},$$

where $\Phi_{l\dots}$ means that all fluxes Φ_l , $l = 1 \dots n$ must be held constant.

Forces and torques are more easily derived from the free enthalpy $G(i_l, x_l, \alpha_l, \dots)$ defined by the Legendre transform

$$G = F - \sum_l i_l \Phi_l,$$

by means of the equations

$$f_{x, \Gamma_k} = - \left. \frac{\partial G}{\partial x_k} \right|_{i_l, \dots}, \quad \tau_\alpha = - \left. \frac{\partial G}{\partial \alpha_k} \right|_{i_l, \dots}.$$

Chapter 9

Magnetic media - Paramagnetism, diamagnetism, ferromagnetism

9.1 The different types of magnetism

In electrostatics, we have seen that matter can be classified into two large families: insulators (or dielectrics) and conductors. In comparison, magnetic materials are classified essentially in three categories depending on their behavior in the presence of an external magnetic field.

- **Ferromagnetism:** Ferromagnetic materials are those that are attracted by an external magnetic field \mathbf{B} , for example, the external field produced by a permanent magnet. Remarkably, these materials remain magnetized even when the external field is removed, becoming themselves permanent magnets. Ferromagnetism is the magnetic property with the largest number of applications in the industry (motors, hard drives, electric generators, etc.).

Examples: Iron (Fe), Cobalt (Co), Nickel (Ni), Iron Oxide (Fe_2O_3).

- **Paramagnetism:** Paramagnetic materials are those that are weakly attracted to the magnetic field \mathbf{B} of a magnet. Unlike ferromagnetic materials, these materials do not exhibit any magnetic activity in the absence of an external magnetic field.

Examples: Aluminum (Al), Calcium (Ca), Sodium (Na).

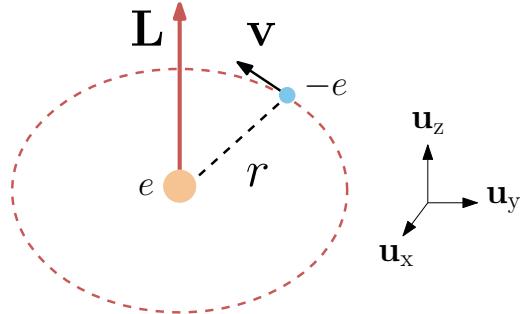
- **Diamagnetism:** Diamagnetic materials are those that are weakly repelled by the magnetic field \mathbf{B} of a magnet. All materials are intrinsically diamagnetic at the atomic level, but their diamagnetic character is generally screened in materials where there is also a paramagnetic or ferromagnetic behavior.

Examples: Bismuth (Bi), Silver (Ag), Copper (Cu), Mercury (Hg), Lead (Pb), water (H_2O).

9.2 Microscopic origin of magnetism

Most of the magnetic properties of matter come from the magnetic dipole moment \mathbf{m} associated to the electrons in atoms. This magnetic moment has two contributions:

- **The orbital magnetic moment:** this can be understood in a simple classical picture. Consider for example the case of hydrogen, where an electron orbits at distance r around a proton.



The angular momentum writes $\mathbf{L} = r\mathbf{u}_r \times m_e \mathbf{v}$, with m_e the electron mass and $r\mathbf{u}_r$ its position in spherical coordinates with respect to the proton. Assuming that the orbit is confined in the Oxy plane, we obtain $\mathbf{L} = rm_e v \mathbf{u}_z$. On the other hand, the motion of the electron around the proton is equivalent to a current loop and therefore a magnetic moment $\mathbf{m}_L = IS\mathbf{u}_z$ with $S = \pi r^2$ the surface area enclosed by the loop and $I = \frac{-ev}{2\pi r}$ the current. This gives

$$\mathbf{m}_L = \frac{-evr}{2} \mathbf{u}_z = \frac{-e}{2m_e} \mathbf{L}. \quad (9.1)$$

The magnetic moment m_L associated with the electron orbit in an atom is therefore proportional to its angular momentum \mathbf{L} . For this reason, it is called the *orbital* magnetic moment. In reality, the angular momentum of the electron is quantized. This is a striking result of quantum mechanics: the angular momentum can only take as possible values an integer multiple of the reduced Planck constant \hbar , that is $L = n\hbar$, where $\hbar = 1.054 \times 10^{-34} \text{ J s}$ and $n \in \mathbb{N}$. In consequence, Eq. (9.1) can be rewritten as:

$$\boxed{\mathbf{m}_L = -\frac{\mu_B}{\hbar} \mathbf{L}}, \quad (9.2)$$

where $\mu_B = \frac{e\hbar}{2m_e} \approx 9 \times 10^{-24} \text{ A m}^2$ is called Bohr's magneton. The orbital magnetic moment is therefore an integer multiple of μ_B , that is, $|\mathbf{m}_L| = n\mu_B$. For this reason, μ_B is called the *quantum* of dipole magnetic moment.

- **The spin magnetic moment:** even in the absence of any orbital motion, the electron has an intrinsic angular momentum \mathbf{S} , called spin, whose projection along an arbitrary axis can take only two possible values, $S = \pm \hbar/2$. This angular momentum does not have any classical interpretation (see Ex. 9.1), it is a purely quantum-mechanical and relativistic effect. The spin contributes to magnetism via the spin magnetic moment

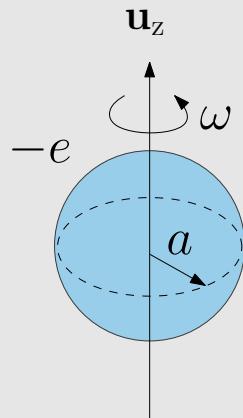
$$\boxed{\mathbf{m}_S = -g \frac{\mu_B}{\hbar} \mathbf{S}}. \quad (9.3)$$

The dimensionless constant $g \approx 2.0023$, called the electron g -factor, is known with an extreme precision. Its departure from 2 can only be explained in the framework of quantum electrodynamics. Note that this intrinsic magnetic moment for the electron has a magnitude of $|m_S| \approx \mu_B$, the Bohr magneton.

- An atom is therefore a complex magnetic system whose total magnetic moment is the result of the several contributions coming from the orbital and spin magnetic moment of the electrons. Note finally that not only the electron has a magnetic moment, but also the nucleons (protons, neutrons), for which the intrinsic magnetic moment m_n is, however, roughly $m_p/m_e \sim 2000$ times smaller than that of an electron, that is $m_n \sim 10^{-3} \mu_B$. The magnetic properties of solids are thus essentially determined by the electrons. Nevertheless, the interaction between the nuclear and electronic spin leads to measurable effects, for example in magnetic resonance this interaction allows the detection, at the atomic scale, of different environments of the nuclei in matter (see Ex. 9.2).

Example 9.1 - Magnetic moment of a rotating charged sphere

Consider a charged sphere of radius a , total charge $-e$ and mass m_e , both uniformly distributed in its volume. The sphere is rotating around the z -axis with an angular velocity ω . The goal of this example is to show that a classical picture of the spin as a self-rotation of the electron is not only inconsistent with experience but also implies rotational speeds larger than the speed of light.



- Write the charge density ρ and the mass density ρ_m in terms of e , m_e and a .
 - Calculate the total angular momentum \mathbf{L} and magnetic moment \mathbf{m} of the rotating sphere. You may use the following result $\int_0^\pi \sin^3 \theta d\theta = 4/3$. Recall that the magnetic moment associated to a current density \mathbf{J} in a volume Ω is given by
- $$m = \frac{1}{2} \iiint_{\Omega} \mathbf{x} \times \mathbf{J}(\mathbf{x}) d^3x.$$
- By writing $\mathbf{m} = g \left(-\frac{e}{2m_e} \right) \mathbf{L}$, determine the g -factor of the self rotating sphere. How does it compare with the experimental value for the electron spin $g \approx 2$?
 - It is known from experiments that every electron has a spin magnetic moment $|\mathbf{m}_S| \approx \mu_B$ where $\mu_B = 9 \times 10^{-24} \text{ A m}^2$ is Bohr's magneton. On the other hand, in example 3.15 in F. Cadiz's textbook, it was shown that if the electron were a uniformly charged sphere, its radius should be $a = r_e = 1.69 \times 10^{-15} \text{ m}$. If the electron spin were due to its self

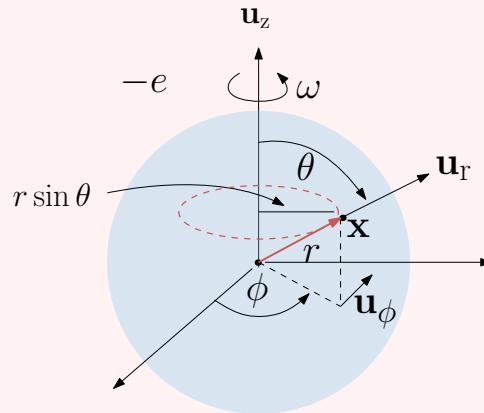
rotation, what would be the required rotation velocity at the *equator* of the electron? Compare with the speed of light $c = 3 \times 10^8 \text{ m s}^{-1}$.

Solution

- a. Since both the charge $-e$ and the mass m_e are uniformly distributed inside the sphere, we have simply

$$\rho = \frac{-e}{\frac{4\pi}{3}a^3}, \quad \rho_m = \frac{m_e}{\frac{4\pi}{3}a^3}.$$

- b. Consider at a certain instant a point $\mathbf{x} = r\mathbf{u}_r$ inside the sphere. The infinitesimal volume $d^3x = r^2dr \sin\theta d\theta d\phi$ around $\mathbf{x} = r\mathbf{u}_r$ describes a circular trajectory of radius $r \sin\theta$ with angular velocity ω .



This elementary volume contains a mass $dm = \rho_m d^3x$ and moves at velocity $\mathbf{v} = \omega r \sin\theta \mathbf{u}_\theta$. Its angular momentum is therefore given by

$$\begin{aligned} d\mathbf{L}(\mathbf{x}) &= \underbrace{\rho_m d^3x}_{dm} r \mathbf{u}_r \times (\underbrace{\omega r \sin\theta \mathbf{u}_\theta}_{\mathbf{v}}) \\ &= \rho_m \omega r^4 \sin^2\theta dr d\theta d\phi (\mathbf{u}_r \times \mathbf{u}_\theta) \end{aligned}$$

and writing $\mathbf{u}_r \times \mathbf{u}_\theta = -\mathbf{u}_\theta = \sin\theta \mathbf{u}_z - \cos\theta(\cos\phi \mathbf{u}_x + \sin\phi \mathbf{u}_y)$, this can be rewritten as

$$\begin{aligned} d\mathbf{L}(\mathbf{x}) &= \rho_m \omega r^4 \sin^3\theta dr d\theta d\phi \mathbf{u}_z - \rho_m \omega r^4 \\ &\quad \times \sin^2\theta \cos\theta dr d\theta d\phi (\cos\phi \mathbf{u}_x + \sin\phi \mathbf{u}_y). \end{aligned}$$

The total angular momentum is obtained by integrating over all the points in the volume Ω of the sphere, after which only the z -component is different from zero

$$\begin{aligned} \mathbf{L} &= \iiint_{\Omega} d\mathbf{L} = \int_0^{2\pi} d\phi \int_0^{\pi} \int_0^a \rho_m \omega r^4 \sin^3\theta dr d\theta \mathbf{u}_z \\ &= 2\pi \rho_m \omega \underbrace{\int_0^{\pi} \sin^3\theta d\theta}_{=\frac{4}{3}} \int_0^a r^4 dr \mathbf{u}_z \\ &= \frac{8\pi \rho_m a^5 \omega}{15} \mathbf{u}_z = \frac{2m_e a^2 \omega}{5} \mathbf{u}_z = \frac{2m_e a^2}{5} \Omega, \end{aligned}$$

where $\boldsymbol{\Omega} = \omega \mathbf{u}_z$. For the magnetic moment, we have

$$\mathbf{m} = \frac{1}{2} \iiint_{\Omega} \mathbf{x} \times \mathbf{J}(\mathbf{x}) d^3x,$$

where the current density at point \mathbf{x} is simply $\mathbf{J}(\mathbf{x}) = \rho(\mathbf{x})\mathbf{v}(\mathbf{x})$ and so,

$$\mathbf{m} = \frac{1}{2} \underbrace{\frac{\rho}{\rho_m}}_{-e/m_e} \underbrace{\iiint_{\Omega} \rho_m \mathbf{x} \times \mathbf{v}(\mathbf{x}) d^3x}_{\mathbf{L}} = \frac{-e}{2m_e} \mathbf{L}.$$

We find the following relationship between the magnetic moment and the angular momentum of the sphere

$$\mathbf{m} = g \frac{-e}{2m_e} \mathbf{L} = \frac{-e}{2m_e} \mathbf{L},$$

so that for a classical rotating sphere $g = 1$, which differs from the experimental value $g \approx 2$ for the electron spin. The spin is a purely quantum mechanical angular momentum that does not come from the electron self rotation.

- c. The magnetic moment of a rotating sphere can be written

$$\mathbf{m} = -\frac{e}{2m_e} \mathbf{L} = -\frac{ea^2}{5} \omega \mathbf{u}_z = -\frac{ea}{5} v_{\text{eq}} \mathbf{u}_z,$$

where $v_{\text{eq}} = a\omega$ is speed at the equator of the sphere. Imposing a to be the electron classical radius r_e , and $|m| = \mu_B$ yields

$$\frac{er_e}{5} v_{\text{eq}} = \mu_B$$

and so,

$$v_{\text{eq}} = \frac{5\mu_B}{er_e} \sim 1.6 \times 10^{11} \text{ m s}^{-1},$$

that is, about 500 times the speed of light! Not only the classical view of the electron spin as a self rotation fails to explain the observed g -factor, but it would also require a rotational speed incompatible with special relativity.

9.3 Dynamics of a magnetic dipole moment in a magnetic field

9.3.1 The Larmor precession

Consider a particle of total angular momentum \mathbf{J} (in an atom, both the orbital and the spin angular momentum contribute). The associated magnetic moment writes, quite generally, as

$$\mathbf{m} = \gamma \mathbf{J},$$

where $\gamma = -g\mu_B/\hbar$ is called the *gyromagnetic ratio*. For example, for the electron spin $\gamma = -g\frac{e}{2m_e}$. In the presence of a uniform magnetic field \mathbf{B} , a torque $\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}$ will act on the particle. The dynamics of its angular momentum \mathbf{J} then writes

$$\frac{d\mathbf{J}}{dt} = \boldsymbol{\tau} = \mathbf{m} \times \mathbf{B},$$

or, in terms of the magnetic dipole moment:

$$\boxed{\frac{d}{dt}\mathbf{m} = \gamma\mathbf{m} \times \mathbf{B}.} \quad (9.4)$$

If the magnetic moment \mathbf{m} is not aligned with the magnetic field \mathbf{B} , Eq. (9.4) describes a precession of \mathbf{m} around the magnetic field axis, called Larmor precession.

- Indeed, taking the scalar product with \mathbf{B} on both sides of Eq. (9.4), we find

$$\frac{d}{dt}(\mathbf{m} \cdot \mathbf{B}) = 0,$$

so that the component of the magnetic moment parallel to the magnetic field is constant.

- Similarly, taking the scalar product with \mathbf{m} at both sides of Eq. (9.4) gives

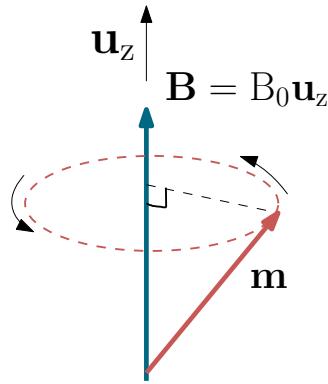
$$\mathbf{m} \cdot \frac{d}{dt}\mathbf{m} = \frac{1}{2} \frac{d}{dt}(|\mathbf{m}|^2) = 0$$

so that $|\mathbf{m}|$ is a constant.

- Assuming a constant magnetic field $\mathbf{B} = B_0 \mathbf{u}_z$ along the z -axis, and defining the Larmor frequency $\omega_L = -\gamma B_0$ we find

$$\begin{aligned} \frac{dm_x(t)}{dt} &= -\omega_L m_y, \\ \frac{dm_y(t)}{dt} &= \omega_L m_x. \end{aligned}$$

The projection of \mathbf{m} in the plane perpendicular to \mathbf{B} describes a circle at the Larmor frequency ω_L .



Remark

The interaction energy between a magnetic moment and a magnetic field writes $W = -\mathbf{m} \cdot \mathbf{B}$, so there is an energetically stable configuration in which the magnetic moment aligns in the direction of the magnetic field. Note, however, that Eq. (9.4) predicts a perpetual movement and it appears that \mathbf{m} never truly aligns with the magnetic field. Of course, this equation describes an isolated system, for which $\mathbf{m} \cdot \mathbf{B} = m_z B_0$ is a constant of motion. For the magnetic dipole to achieve the position $\mathbf{m}_{\text{eq}}(\mathbf{B})$ that minimizes the energy, corresponding to thermodynamical equilibrium, energy exchanges with a reservoir must be taken into account. Whenever the magnetic moment is not in its equilibrium position, we can add a term in Eq. (9.4) to allow for an exponential return to equilibrium with a characteristic time T_r , also called spin-lattice relaxation time

$$\boxed{\frac{d\mathbf{m}}{dt} = \gamma \mathbf{m} \times \mathbf{B} - \frac{\mathbf{m} - \mathbf{m}_{\text{eq}}(\mathbf{B})}{T_r}.} \quad (9.5)$$

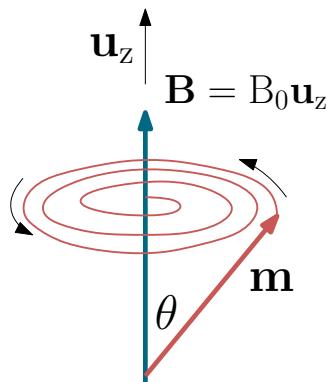
For example, for a field $\mathbf{B} = B_0 \mathbf{u}_z$ for which the equilibrium position is $\mathbf{m}_{\text{eq}}(\mathbf{B}) = m_s \mathbf{u}_z$, projecting Eq. (9.5) on the z -axis gives

$$\frac{dm_z(t)}{dt} = -\frac{m_z(t) - m_s}{T_1}.$$

For an initial condition $m_z(0) = m_s \cos \theta$, we obtain the following solution for the z -component of the magnetic moment

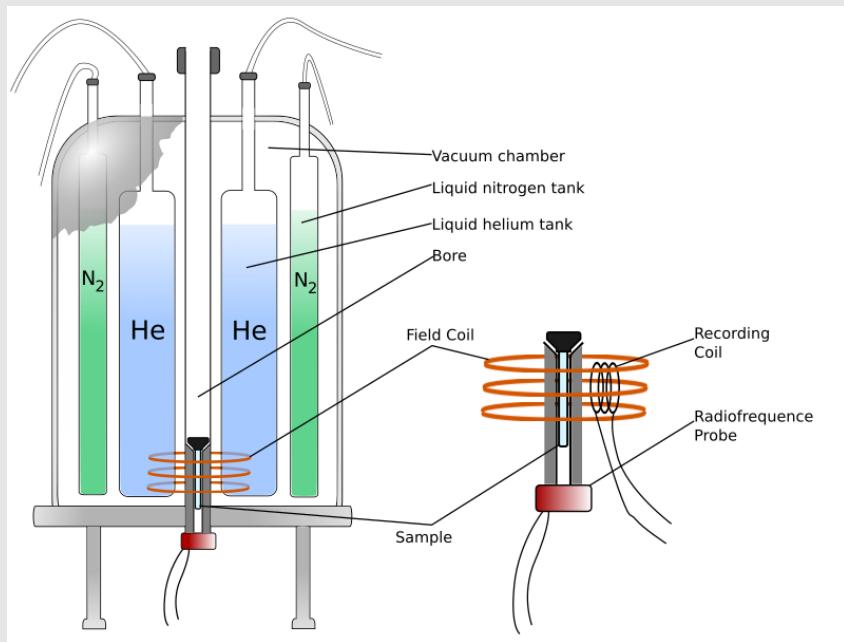
$$m_z(t) = m_s \left(1 - e^{-t/T_1} \right) + m_s \cos \theta e^{-t/T_1}.$$

In addition to the Larmor precession of the components perpendicular to \mathbf{B} , we see that a magnetic moment reaches its equilibrium position by describing a spiral.



In a magnetic medium that does not carry a permanent dipole moment, thermal disorder is responsible for a random orientation of the atomic magnetic moments. If at $t = 0$ we apply a magnetic field to this sample, it will become magnetized, since all the magnetic moments will align with the magnetic field after a time T_1 .

Example 9.2 - Applications: Electron-spin resonance and Nuclear magnetic resonance



Electron-spin resonance (ESR) denotes the phenomenon of selective absorption of weak sinusoidal electromagnetic radiation in the microwave region by unpaired electrons in the atomic structure of paramagnetic materials that simultaneously are subjected to a constant, strong magnetic field.

When a strong stationary magnetic field is exerted on materials containing unpaired electrons, the latter behave like tiny magnets because of their spin. The dipole moments partially align with the strong external field, and precess in the field (Larmor precession). Absorption of energy from the weak oscillating magnetic field occurs when the frequency of the weak oscillating field corresponds to the natural precession frequency of the dipole moments. This is the resonance phenomenon. The amount of absorbed radiation from the weak field then depends on the microwave frequency of the weak field and the strength of the strong stationary field. Hence, an electron paramagnetic resonance spectrum can be measured by varying one parameter while keeping the other fixed, that is, typically, by recording the absorbed energy of the microwave radiation versus the applied strength of magnetic field.

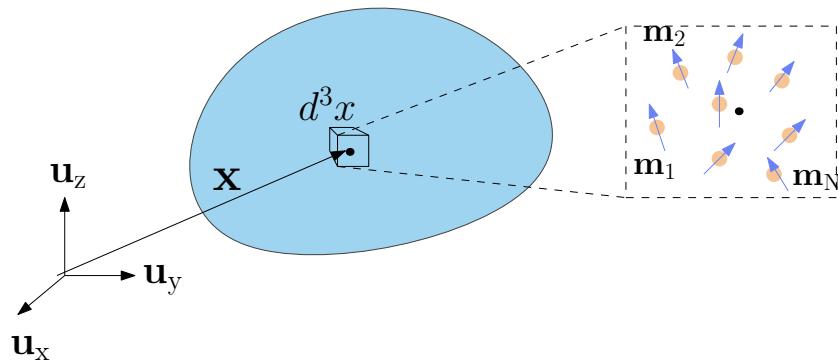
Measuring the Larmor frequency for electron spins is a very useful spectroscopic technique capable of identifying paramagnetic substances. It permits investigations on the nature of chemical bonds within molecules by identifying unpaired electrons and their interaction with the immediate surroundings. It is used in biology, chemistry and physics and has medical applications as well.

There is also a characteristic Larmor frequency for nuclear spins. The characteristic frequencies associated with electron spin are employed in electron spin resonance (ESR) experiments, and those associated with the nuclear spin in nuclear magnetic resonance (NMR) experiments. In the latter case, the atomic electron cloud tends to screen the effective magnetic field seen by the nuclei and as a result, the nuclear spin resonance depends on the electron density distribution in the corresponding molecular orbitals. Nuclear magnetic resonance spectroscopy is widely used to determine the structure of organic molecules in solution and study molecular physics, crystals as well as non-crystalline materials. NMR is also routinely used in advanced medical imaging techniques, such as in magnetic resonance imaging (MRI).

Suggested Exercise: Ex. 9.2 'Magnetic spin resonance' in the exercise book.

9.4 Magnetization of matter

Since every magnetic dipole moment \mathbf{m} generates a magnetic field and in turn interacts with external magnetic fields, every atom with a non-zero magnetic moment behaves as a tiny magnet. A material of volume Ω may exhibit a macroscopic magnetic character if, locally, there exists a certain order of the individual magnetic dipole moments, as shown in the figure below where the magnetic moments seem to preferentially align along the z -axis. If such an order does not exist, the average magnetic moment around any point in Ω will be negligible, in which case the material is said to be demagnetized or simply non-magnetic.



Definition 9.1: Magnetization

Let Ω be a certain volume of matter and d^3x a *mesoscopic* volume around a point $\mathbf{x} \in \Omega$ containing a significant amount N of atoms. The magnetization \mathbf{M} represents a spatially averaged magnetic moment density such that, at every $\mathbf{x} \in \Omega$,

$$\mathbf{M}(\mathbf{x})d^3x = \sum_{n \in d^3x} \mathbf{m}_n, \quad (9.6)$$

where the index $n = 1, 2, \dots, N$ counts the atoms contained in d^3x .

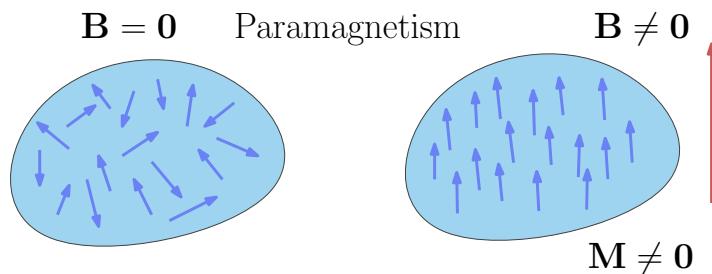
9.4.1 The microscopic origin of diamagnetism

Most insulating materials present a very weak magnetism, and are generally diamagnetic. This comes from the fact that an insulator is an assemblage of atoms with saturated electronic shells, for which the orbital and spin magnetic moments compensate each other. In this case, the only source of magnetism is a quantum effect¹ that occurs through the modification of the atomic orbitals in the presence of an external magnetic field, called diamagnetism.

Essentially, a diamagnetic material responds to an external magnetic field by the appearance of a magnetization \mathbf{M} that tends to cancel out the total field inside it. This manifests macroscopically by a repulsive force whose most spectacular application is magnetic levitation. Superconductors, from a certain point of view, can be considered as perfectly diamagnetic, since the magnetic field is completely canceled inside them (Meissner effect).

9.4.2 Microscopic origin of Paramagnetism and Ferromagnetism

In metals, free electrons are responsible for what is called the Pauli spin paramagnetism. Since each electron has a spin, under the presence of an external magnetic field, one spin configuration becomes energetically favorable, inducing at thermal equilibrium a small magnetization $\mathbf{M} \neq \mathbf{0}$ aligned with the external magnetic field. A stronger form of paramagnetism, called the Curie paramagnetism, exists in solids having ions with unsaturated electronic shells with a net magnetic moment. These magnetic moments align as well in the presence of an external magnetic field. Regardless of the type of paramagnetism, when the external magnetic field is turned off, magnetization disappears due to thermal disorder.

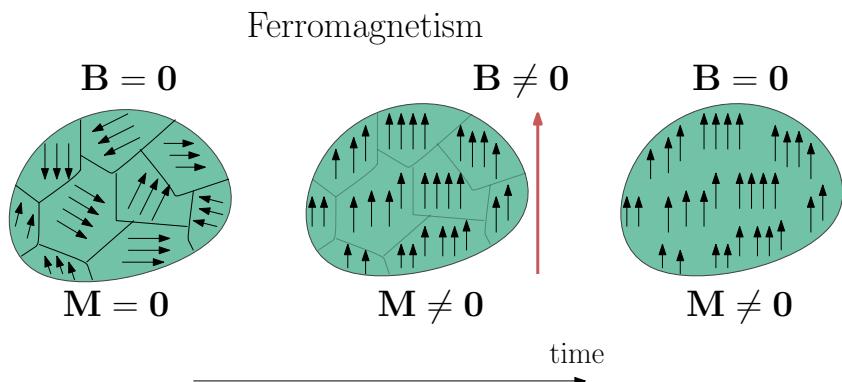


The Curie paramagnetism increases with decreasing temperature, until a critical temperature T_C called the Curie temperature is achieved (typically of several hundred kelvins), below which the no longer negligible interaction between neighboring magnetic moments is responsible for the emergence of a spontaneous magnetization: even in the absence of a magnetic field, magnetic moments may be locally aligned resulting in the appearance of magnetic domains (Weiss domains), generally separated by crystalline defects. The material has become ferromagnetic.

These magnetic domains will remain aligned due to their mutual interactions, thus generating a *remanent magnetization*, i.e., a non zero magnetization that persists in the absence of an external

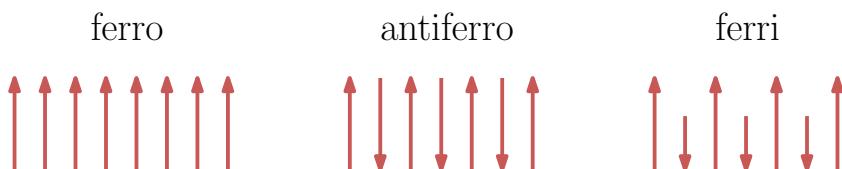
¹Although a semi-classical, qualitative model due to Langevin allows us to understand diamagnetism in terms of Faraday's law of induction.

magnetic field. In conclusion, the paramagnetic or ferromagnetic behavior of a material is a matter of energy scales. It is therefore dependent on temperature.



At first glance, one may think that ferromagnetic order arises from dipole-dipole interactions between neighboring atoms. However, this interaction predicts critical temperatures of the order of $T_C = 0.1\text{ K}$ (to compare for example with the case of iron for which $T_C = 1043\text{ K}$). In fact, the interaction responsible for the magnetic order has a completely quantum origin: it results from a combination of Pauli's exclusion principle and the Coulomb interaction between electrons. Essentially, the electrostatic repulsion between two electrons depends on the relative orientation of their spins.

Other forms of collective magnetism exist: in an antiferromagnetic material, the magnetic dipole moments are anti-parallel between first neighbors. This configuration is possible for example if a crystal lattice can be simply divided into two equivalent sub-lattices. If the two sub-lattices are made of atoms of the same magnetic moment, the net magnetization is zero. However, if the two sub-lattices have atoms with different magnetic moments, a net magnetization appears under an antiferromagnetic arrangement, and these materials are called ferrimagnetics. This is the case of ferrites (Fe_2O_3 , Fe_3O_4).



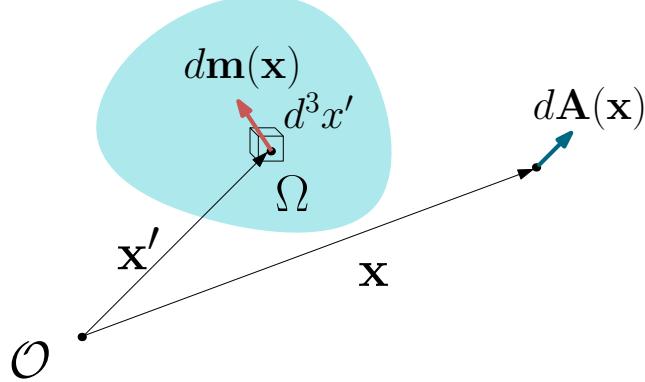
9.4.3 From electronics to spintronics

The electron spin, responsible for ferromagnetism, is nowadays used as a means of writing and reading information in hard drives. A hard drive is made of tiny ferromagnetic domains whose magnetization orientation can be independently set in order to represent ones and zeros (data bits). The data is read electrically by exploiting a phenomenon called giant magnetoresistance: due to the electron spin, the resistance between two ferromagnetic metals depends strongly on the relative orientation of their magnetizations. The discovery of the giant magnetoresistance in 1988 by Albert Fert and Peter Grünberg was awarded by the Nobel prize in 2007. There is nowadays a large research effort trying to exploit the electron spin to create new optoelectronic devices with better performances. This emergent nanotechnology is called spintronics.



9.5 Magnetic field generated by a magnetized material

Consider a volume Ω characterized by a magnetic dipole density \mathbf{M} , so that a volume element d^3x' around \mathbf{x}' has a total dipole moment $d\mathbf{m}(\mathbf{x}') = \mathbf{M}(\mathbf{x}')d^3x'$.



The infinitesimal vector potential at a point \mathbf{x} due to this infinitesimal magnetic dipole moment writes

$$d\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \frac{d\mathbf{m}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = \frac{\mu_0}{4\pi} \frac{\mathbf{M}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'$$

and the total potential at \mathbf{x} is obtained by integrating over Ω :

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{M}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x'.$$

Now, we use the following identity:

$$\mathbf{M}(\mathbf{x}') \times \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = \mathbf{M}(\mathbf{x}') \times \nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{\nabla' \times \mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} - \nabla' \times \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right),$$

where ∇' denotes differentiation with respect to the coordinates of \mathbf{x}' , so that

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\nabla' \times \mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' - \frac{\mu_0}{4\pi} \iiint_{\Omega} \nabla' \times \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right) d^3x'.$$

Finally, transforming the second integral into a surface integral (detailed below):

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\nabla' \times \mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{\mu_0}{4\pi} \oint_{\partial\Omega} \frac{\mathbf{M}(\mathbf{x}') \times \mathbf{n}(\mathbf{x}') dS(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (9.7)$$

We see that (9.7) looks like the general solution to Poisson's equation for the vector potential in magnetostatics generated by a current distribution of volume density $\mathbf{J}(\mathbf{x}')$ and surface density $\mathbf{j}_s(\mathbf{x}')$:

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{\mu_0}{4\pi} \oint_{\partial\Omega} \frac{\mathbf{j}_s(\mathbf{x}') dS(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}.$$

This will allow us to identify *magnetization currents* below.

Now we complete the demonstration of (9.7) with the detail of the transformation into a surface integral of the last term of the right-hand side. For this demonstration, consider a constant unit vector \mathbf{u} , for example in a Cartesian basis. We will show that the transformation into a surface integral can be done for the projection on \mathbf{u} of \mathbf{A}_2 , the second term on the right hand side of (9.7). Since \mathbf{u} is a constant vector, the scalar product with \mathbf{u} can be performed under the integral:

$$\mathbf{A}_2 \cdot \mathbf{u} = - \iiint_{\Omega} \frac{\mu_0}{4\pi} \left[\nabla' \times \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right) \right] \cdot \mathbf{u} d^3x'$$

Now we use a general differentiation rule of vector calculus: $(\nabla \times \mathbf{a}(\mathbf{x})) \cdot \mathbf{u} = \nabla \cdot (\mathbf{a}(\mathbf{x}) \times \mathbf{u})$, with $\mathbf{a}(\mathbf{x}') \equiv \mathbf{M}(\mathbf{x}')/|\mathbf{x} - \mathbf{x}'|$, and we use Ostrogradsky's theorem to transform the resulting volume integral of a divergence into a surface integral:

$$\mathbf{A}_2 \cdot \mathbf{u} = - \iiint_{\Omega} \frac{\mu_0}{4\pi} \nabla' \cdot \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \times \mathbf{u} \right) d^3x' = - \iint_S \frac{\mu_0}{4\pi} \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \times \mathbf{u} \right) \cdot \mathbf{n}(\mathbf{x}') dS(\mathbf{x}').$$

A permutation of \mathbf{u} and \mathbf{n} in the triple product, and correspondingly a change of sign, leads to the result:

$$\mathbf{A}_2 \cdot \mathbf{u} = \left(\iint_S \frac{\mu_0}{4\pi} \frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \times \mathbf{n}(\mathbf{x}') dS(\mathbf{x}') \right) \cdot \mathbf{u}.$$

Since this is true for any unit vector of a Cartesian basis and thus for all components of the vector \mathbf{A}_2 , we have

$$\mathbf{A}_2 = - \iiint_{\Omega} \frac{\mu_0}{4\pi} \left[\nabla' \times \left(\frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right) \right] = \iint_S \frac{\mu_0}{4\pi} \frac{\mathbf{M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \times \mathbf{n}(\mathbf{x}') dS(\mathbf{x}'),$$

which completes the demonstration of (9.7).

9.5.1 The magnetization current densities

Recalling that the vector potential of a current distribution in a volume Ω is given by

$$\mathbf{A}(\mathbf{x}) = \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x',$$

we see from Eq. (9.7) that a magnetized medium is equivalent to a medium having a current density in Ω :

$$\mathbf{J}_M(\mathbf{x}) = \nabla \times \mathbf{M}(\mathbf{x}) \quad \mathbf{x} \in \Omega,$$

and a surface current density

$$\mathbf{J}_{S,M}(\mathbf{x}) = \mathbf{M}(\mathbf{x}) \times \mathbf{n}(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega,$$

so that

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}_M(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{\mu_0}{4\pi} \iint_{\partial\Omega} \frac{\mathbf{J}_{S,M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} dS(\mathbf{x}').$$

Definition 9.2: Magnetization currents

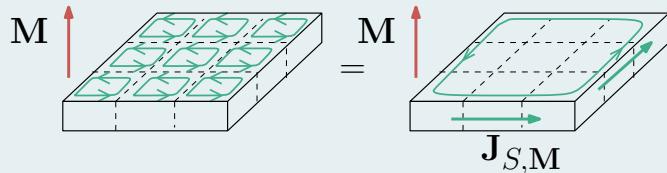
The magnetic field generated by a material Ω of magnetization \mathbf{M} is equivalent to that generated by a volume carrying a current density \mathbf{J}_M in Ω and $\mathbf{J}_{S,M}$ at its boundary $\partial\Omega$, with

$$\mathbf{J}_M(\mathbf{x}) = \nabla \times \mathbf{M}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (9.8)$$

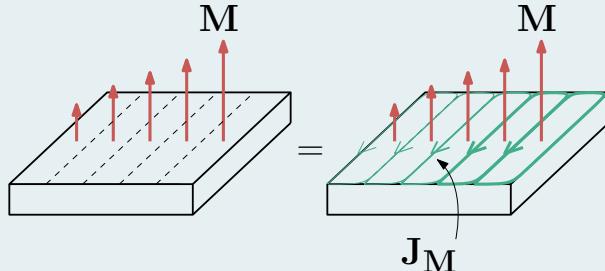
and

$$\mathbf{J}_{S,M}(\mathbf{x}) = \mathbf{M}(\mathbf{x}) \times \mathbf{n}(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega, \quad (9.9)$$

where \mathbf{n} is normal to $\partial\Omega$. The interpretation of these magnetization currents, also called Amperian currents, is the following: a medium with uniform magnetization \mathbf{M} can be thought of being made of microscopic current loops, all equal and oriented in the same way as shown in the figure below. In the bulk, these currents cancel each other at any point. The result is a net current flowing only on the surface of the magnetized medium, of density $\mathbf{J}_{S,M} = \mathbf{M} \times \mathbf{n}$.



If a nonuniform magnetization is present, the compensation between these current loops is not perfect and a volumetric current develops.



Note that

$$\nabla \cdot \mathbf{J}_M(\mathbf{x}) = 0,$$

as it should be for a stationary current, and that

$$\mathbf{n}(\mathbf{x}) \cdot \mathbf{J}_{S,M}(\mathbf{x}) = 0,$$

as should happen for a current located on $\partial\Omega$. Then, the effect of an external magnetic field on a magnetic medium can be summarized as the generation of a stationary current located within the medium.

Let us finally note that, while in a dielectric the polarization charge density represents a true charge density resulting from the spatial separation between positive and negative charges, the magnetization current density represents a fictitious current in the sense that it is not associated with the motion of

charges: the spin magnetic moment, for example, does not come from any electrical current.

9.6 Magnetostatics in a magnetized medium

Now we will see how to rewrite the fundamental laws of the magnetic field inside a magnetized medium. In the most general case, at any point in space there can be a *free* current density (also called a *conduction* current density), $\mathbf{J}_{\text{free}}(\mathbf{x})$, and a magnetization current density $\mathbf{J}_M(\mathbf{x}) = \nabla \times \mathbf{M}(\mathbf{x})$. Both currents densities must be included when writing Ampère's theorem for the spatially averaged magnetic field:

$$\nabla \times \langle \mathbf{B} \rangle(\mathbf{x}) = \mu_0 (\mathbf{J}_{\text{free}}(\mathbf{x}) + \nabla \times \mathbf{M}(\mathbf{x})).$$

In the following, we will simply write \mathbf{B} instead of $\langle \mathbf{B} \rangle$, but it must be recalled that we are dealing with spatially averaged quantities so that a medium can be characterized by a smooth magnetization \mathbf{M} . Equivalently, Ampère's law can be written as

$$\nabla \times \left(\frac{\mathbf{B}(\mathbf{x})}{\mu_0} - \mathbf{M}(\mathbf{x}) \right) = \mathbf{J}_{\text{free}}(\mathbf{x}).$$

This way of writing Ampère's law has the advantage of keeping on the right hand side only the free, real current density.

Definition 9.3: Magnetic excitation and Ampère's theorem

We define the vector field \mathbf{H} , called magnetic excitation, by:

$$\mathbf{H}(\mathbf{x}) = \frac{\mathbf{B}(\mathbf{x})}{\mu_0} - \mathbf{M}(\mathbf{x}), \quad (9.10)$$

which satisfies Ampère's theorem in matter where only the free current density appears explicitly:

$$\nabla \times \mathbf{H}(\mathbf{x}) = \mathbf{J}_{\text{free}}(\mathbf{x}). \quad (9.11)$$

The entire effect of the magnetized medium is thus contained in \mathbf{H} through the magnetic moment density \mathbf{M} . The S.I. unit for the excitation field \mathbf{H} is the same as for the magnetization: $[\text{A m}^{-1}]$, which represents a current per unit length.

The fundamental equations of magnetostatics in a magnetized medium are then

$$\begin{aligned} \nabla \cdot \mathbf{B}(\mathbf{x}) &= 0, \\ \nabla \times \mathbf{H}(\mathbf{x}) &= \mathbf{J}_{\text{free}}(\mathbf{x}). \end{aligned}$$

Note the total analogy with the fundamental laws of electrostatics in a dielectric medium

$$\begin{aligned} \nabla \cdot \mathbf{D}(\mathbf{x}) &= \rho_{\text{free}}(\mathbf{x}), \\ \nabla \times \mathbf{E}(\mathbf{x}) &= \mathbf{0}, \end{aligned}$$

where $\rho_{\text{free}}(\mathbf{x})$ is the free charge density. In linear, homogeneous and isotropic dielectrics, $\mathbf{D}(\mathbf{x}) = \epsilon \mathbf{E}(\mathbf{x})$, so that the two fundamental equations of electrostatics can be written as:

$$\nabla \cdot \mathbf{E}(\mathbf{x}) = \frac{\rho_{\text{free}}(\mathbf{x})}{\epsilon},$$

$$\nabla \times \mathbf{E}(\mathbf{x}) = 0,$$

where the main difference with the laws in vacuum is the permittivity of the medium $\epsilon > \epsilon_0$. This implies that the polarization of a dielectric medium decreases the magnitude of the electric field generated by free charges due to screening. The relation between \mathbf{D} and \mathbf{E} , an equation of state for the dielectric medium, closes Maxwell's equation of electrostatics in a dielectric medium.

We will see that a similar treatment can be performed for magnetic materials, by writing an equation of state (or constitutive relation for the magnetic medium, that is, a relation between the magnetic field \mathbf{B} and the magnetic excitation \mathbf{H}). This relation will close Maxwell's equations for magnetostatics in a magnetic medium.

9.7 Constitutive relations of a magnetic material

In a linear, homogeneous and isotropic magnetic material, depending on whether the response is *paramagnetic* or *diamagnetic*, the magnetic field generated by a free current can increase or decrease with respect to the value it would have in vacuum, so that everything happens as if the permeability in the medium, μ , was renormalized with respect to μ_0 . We will see, in addition, that the response of a *ferromagnetic* medium depends on the history of magnetization history (a phenomenon called hysteresis), and so it is not possible in that case to define an equivalent permeability μ without knowing the history of magnetization.

Definition 9.4: Magnetic susceptibility and permeability

In a medium which is homogeneous and isotropic, experience shows that in the static regime the material response is such that

$$\mathbf{M}(\mathbf{x}) = \chi_m(\mathbf{H})\mathbf{H}(\mathbf{x}),$$

where the dimensionless quantity $\chi_m(\mathbf{H})$ is the *magnetic susceptibility*. We see then that

$$\mathbf{H}(\mathbf{x}) = \frac{\mathbf{B}(\mathbf{x})}{\mu_0} - \chi_m(\mathbf{H}(\mathbf{x}))\mathbf{H}(\mathbf{x})$$

and so,

$$\mathbf{B} = \mu_0(1 + \chi_m(\mathbf{H}))\mathbf{H} = \mu(\mathbf{H})\mathbf{H}, \quad (9.12)$$

where $\mu(\mathbf{H})$ is the *magnetic permeability* of the material. In addition, if the material is *linear*, its magnetic permeability does not depend on \mathbf{H} , that is, in a linear homogeneous and isotropic material, the magnetic field is proportional to the magnetic excitation.

With this, the fundamental laws of magnetostatics in a linear, homogeneous and isotropic medium can be written as follows:

$$\nabla \cdot \mathbf{B}(\mathbf{x}) = 0,$$

$$\nabla \times \mathbf{B}(\mathbf{x}) = \mu \mathbf{J}(\mathbf{x}),$$

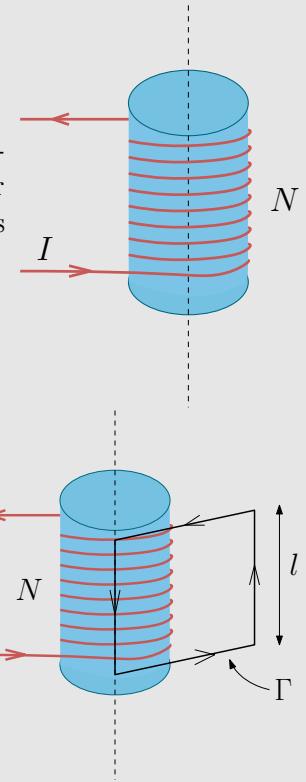
that is, as in a vacuum, but renormalizing the permeability $\mu \neq \mu_0$.

For a diamagnetic or paramagnetic material, the susceptibility does not depend on \mathbf{H} : $\chi_m(\mathbf{H}) = \chi_m$, and the material response is linear. Typically, in a diamagnetic medium, the susceptibility is $\chi_m \sim -10^{-5} < 0$ (so that $\mu < \mu_0$) and independent of temperature since diamagnetism is a property of the atomic orbital occupied by the electrons, which is essentially the one corresponding to the lowest energy configuration. This diamagnetism is negligible, for example, with respect to that of superconductors ($\chi \sim -1$).

In a paramagnetic medium typically $0 \lesssim \sim 10^{-5} \lesssim \chi_m - 10^{-3}$ (and $\mu > \mu_0$). For a ferromagnetic material, on the other hand, the susceptibility depends strongly on \mathbf{H} : $\chi_m(\mathbf{H})$ can vary between 50 and 10^6 depending on the value of \mathbf{H} . The response of a ferromagnet is therefore nonlinear and one may achieve very high values for the permeability ($\mu(\mathbf{H}) \gg \mu_0$).

Example 9.3 - The magnetic excitation vector as a control parameter

The interest of the vector \mathbf{H} resides in the fact that one can usually directly control its strength through an external current. In fact, consider a solenoid of N turns that carries a free current I_{ext} , in a wire that is wrapped around a soft iron core, as shown in the following figure.



Assuming an infinitely long solenoid^a, the symmetry of the current density \mathbf{J}_{ext} on the surface of the cylindrical solenoid implies a uniform \mathbf{H} field inside the solenoid, along the direction of its axis. In that case, we may use Ampère's theorem (9.11) integrated over the closed, squared planar contour Γ of side l shown in the figure below.

^athat is, the length is much larger than the radius and we can safely neglect edge effects.

We have

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{l} = Hl = \frac{Nl}{L} I_{\text{ext}},$$

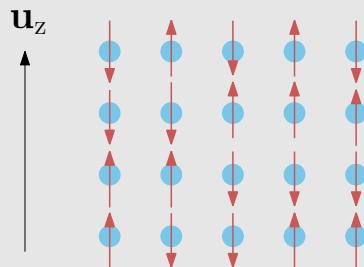
where L is the total length of the coil. In terms of the density $n = N/L$ of turns, we obtain

$$H = nI_{\text{ext}},$$

so that in a magnetic material, it is the magnetic excitation field \mathbf{H} , and not directly \mathbf{B} , which is controlled by an external current. The magnetic field \mathbf{B} will indeed depend on the magnetic permeability $\mu\mathbf{H}$ of the soft iron core.

Example 9.4 - Curie paramagnetism of localized spins

Consider an ensemble of localized, non-interacting spins. In the absence of magnetic field, the spins are oriented either parallel ($\mathbf{m} = +\mu_B \mathbf{u}_z$) or antiparallel ($\mathbf{m} = -\mu_B \mathbf{u}_z$) to the z -axis with equal probability, so that $\mathbf{M} = 0$.



- a. Now a magnetic field $\mathbf{B} = B_0 \mathbf{u}_z$ is applied. What are the energies of the two possible configurations ($\mathbf{m} = \pm \mu_B \mathbf{u}_z$) for a spin in this magnetic field? Which state is energetically favorable?
- b. The spins will have a tendency to adopt the lowest energy configuration, in a competition with thermal agitation. This can be taken into account by assuming that the probability P for a spin to occupy a state of energy E is proportional to $e^{-E/k_B T}$, where T is the temperature and k_B the Boltzmann constant. Calculate the averaged magnetic moment of a single spin. What happens in the two limiting cases $\mu_B B_0 \gg k_B T$ and $\mu_B B_0 \ll k_B T$? Which one is the relevant limit at room temperature for a magnetic field of the order of 1 Tesla?
- c. Let n be the number of spins per unit volume and suppose $\mu_B B_0 \ll k_B T$. Show that the magnetic susceptibility obeys Curie's law:

$$\chi_m \propto \frac{1}{T}$$

and find the proportionality constant. What is the order of magnitude of the magnetic susceptibility for $n = 1 \times 10^{21} \text{ cm}^{-3}$ at room temperature?

Solution

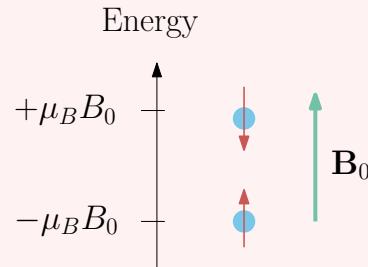
- a. The energy of a magnetic dipole \mathbf{m} in a magnetic field \mathbf{B} writes $E = -\mathbf{m} \cdot \mathbf{B}$. Therefore,

the two possible orientations of a given spin have energies given by

$$E(\mathbf{m} = +\mu_B \mathbf{u}_z) = -\mu_B B_0 = E_1,$$

$$E(\mathbf{m} = -\mu_B \mathbf{u}_z) = +\mu_B B_0 = E_2 > E_1.$$

When the spin is parallel to the magnetic field, the energy is lower by $2\mu_B B_0$, and therefore more energetically favorable than when the spin is antiparallel to the field.



b. the probability for the spin to occupy the state $\pm\mu_B \mathbf{u}_z$ is given by

$$P_{\pm} = C e^{\pm W_0}, \quad \text{with } W_0 = \frac{\mu_B B_0}{k_B T},$$

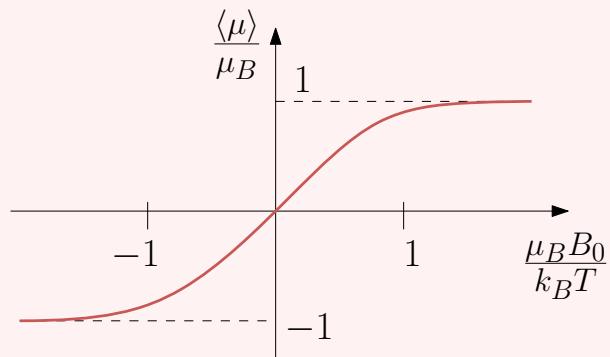
where C is a constant that can be determined by considering that the probability for the spin to occupy either $+\mu_B \mathbf{u}_z$ or $-\mu_B \mathbf{u}_z$ is simply equal to one

$$P_+ + P_- = 1 \quad \rightarrow \quad C = \frac{1}{e^{W_0} + e^{-W_0}}.$$

The average magnetic moment $\langle m \rangle$ along the z -axis of a single spin is therefore

$$\begin{aligned} \langle m \rangle &= +\mu_B P_+ - \mu_B P_- = \mu_B (P_+ - P_-) \\ &= \mu_B \frac{e^{+W_0} - e^{-W_0}}{e^{W_0} + e^{-W_0}} = \mu_B \tanh(W_0). \end{aligned}$$

We see that the spin aligns itself with the magnetic field ($\langle m \rangle \sim \mu_B$) when $W_0 \gg 1$, i.e., when the magnetic interaction energy $\mu_B B_0$ is large compared with the available thermal energy $k_B T$.



In contrast, in the regime where $\mu_B B_0 \ll k_B T$, then $\tanh\left(\frac{\mu_B B_0}{k_B T}\right) \approx \frac{\mu_B B_0}{k_B T}$ so that

$$\langle m \rangle \approx \frac{\mu_B^2 B_0}{k_B T} \ll \mu_B.$$

At room temperature, $k_B T \approx 26 \text{ meV}$ and for 1 Tesla, $\mu_B B_0 = 56 \mu\text{eV}$ so that $\mu_B B_0 / k_B T \sim 10^{-3}$.

c. If n is the density of spins, then the magnetization \mathbf{M} is given by

$$\mathbf{M} = n \langle m \rangle \mathbf{u}_z = n \mu_B \tanh\left(\frac{\mu_B B_0}{k_B T}\right) \mathbf{u}_z.$$

In the limit $\mu_B B_0 \ll k_B T$

$$\mathbf{M} \approx n \frac{\mu_B^2 B_0}{k_B T} \mathbf{u}_z = n \frac{\mu_B^2}{k_B T} \mathbf{B}$$

and since $\mathbf{M} = \frac{\mathbf{B}}{\mu_0} \frac{\chi_m}{(1 + \chi_m)}$, for a weak susceptibility, we find

$$\mathbf{M} \approx \frac{\mathbf{B}}{\mu_0} \chi_m = n \frac{\mu_B^2}{k_B T} \mathbf{B},$$

so finally

$$\chi_m = n \mu_0 \frac{\mu_B^2}{k_B T}.$$

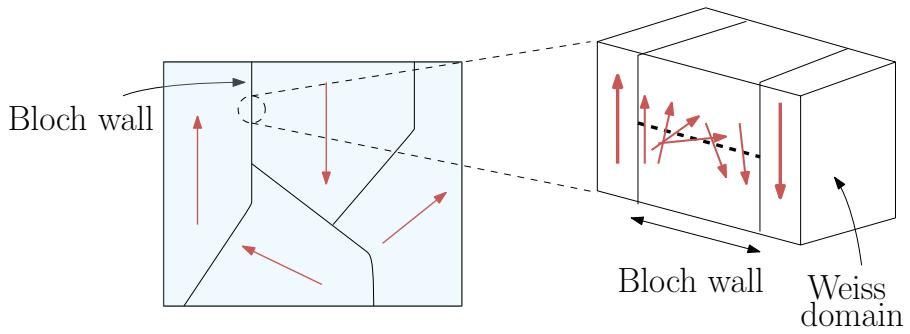
For $n = 1 \times 10^{21} \text{ cm}^{-3}$ and $k_B T \sim 26 \text{ meV}$, we find

$$\chi_m \sim 10^{-5}.$$

9.8 Hysteresis

The increase in the magnitude of the magnetic field inside a ferromagnetic material can be considerable, up to 10^6 times larger than that of the applied external field. The behavior of a ferromagnet is **nonlinear** and leads to the most spectacular of magnetic effects. Remarkably, the relationship between the magnetization \mathbf{M} and \mathbf{B} (or equivalently, between \mathbf{M} and \mathbf{H}) is not single valued, it depends on the history of the material. This phenomenon is known as **hysteresis**.

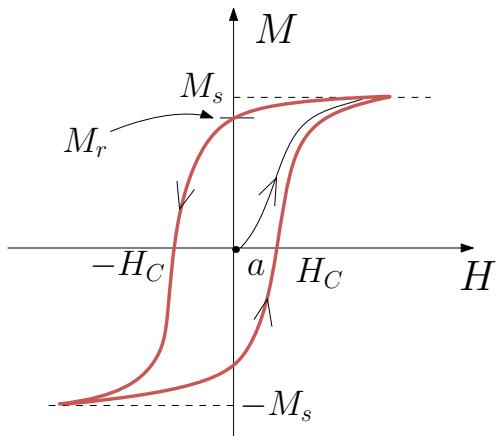
Consider an initially demagnetized ferromagnetic material. Because of the presence of defects, it is composed of microscopically small domains (sizes between 10 and 100 μm), also called Weiss domains, in which all the atomic magnetic moments point in the same direction, and so a non-zero, constant magnetization is present in every Weiss domain. Between two domains, the magnetization can vary, and in the very thin region (of typical width of 0.1 μm) that separates them, also called Bloch wall, the magnetization is reversed continuously.



The size of the domains results from a compromise between the bulk magnetic energy stocked inside a Weiss domain ($U_{\text{dom}} = \iiint_{\Omega} \frac{B^2}{2\mu} d^3x$) and the energy cost of creating a wall, that is, to reverse the magnetization. Minimization of energy requires a fragmentation into volumes with randomly oriented magnetization so as to minimize the magnetic field inside the material.

Now, when a weak magnetic field \mathbf{H} is applied, there is a displacement of the walls so as to favor the fraction of domains whose magnetization is parallel to that of the external field. The magnetization \mathbf{M} starts to increase in a linear, reversible way. But if the field \mathbf{H} is large enough, then sudden displacements of walls are produced followed by the disappearance of domains. These processes are generally irreversible after which the magnetization keeps increasing by a reversible rotation of the magnetic moments and reaches a saturation value M_s when all the magnetic moments are parallel to \mathbf{H} . If now the field is decreased, the magnetization decreases as well but at a slower rate, and at $\mathbf{H} = \mathbf{0}$ there is a remanent (i.e. nonzero) magnetization M_r . The ferromagnet is now permanently magnetized, it becomes a permanent magnet.

In order to cancel this magnetization a field in the opposite direction must be applied. Magnetization then vanishes $M = 0$ for a value of $H = -H_c$ called the coercive field. If the sample is exposed to the cycle $0 \rightarrow H \rightarrow -H \rightarrow H$, one obtains a hysteresis cycle. The figure below shows the variation of the magnetization in the presence of an external magnetic field \mathbf{H} , starting from a situation in which the material is initially demagnetized (point a). The hysteresis cycle is characterized by the values of M_r and H_c , essential parameters when choosing ferromagnetic materials.



There are two main categories of ferromagnetic materials:

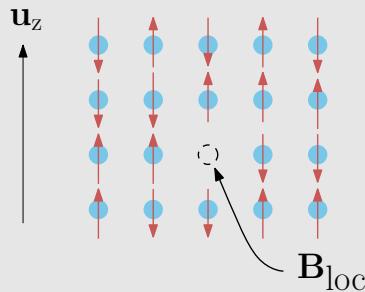
- **Soft ferromagnetics** present a weak coercive field, a weak remanent magnetization and a very large magnetic permeability at saturation, allowing a concentration of the magnetic flux inside them. They are used to fabricate electromagnets, since they are easy to demagnetize, and are also widely found in electric transformers. An example is a Ni-Fe alloy called Mu-metal, exhibiting a very high permeability of $\mu \sim 10^5 \mu_0$, used for shielding electronic devices against static or slowly varying magnetic fields (see Ex. 9.7 in F. Cadiz's textbook). In the absence of magnetic excitation, they do not create a magnetic field in their surrounding.

- **Hard ferromagnetics**, also called magnets, have instead a large coercitive field, and so they are hard to demagnetize. These materials are ideal for the fabrication of permanent magnets, such as those used in electric motors or generators. An example are the rare earth NdFeB magnets, the strongest commercially available, and for which the coercitive field is of the order of $H_C = 10^6 \text{ A m}^{-1}$. The magnetic field at their surface can reach typically one tesla.

Example 9.5 - Transition between paramagnetism and ferromagnetism

It is well known that Curie's law $\chi_m \propto 1/T$ is only valid at high temperature, since when the latter is low enough, the interaction between magnetic moments is responsible for the appearance of a collective magnetization, i.e., ferromagnetism. A simple model due to Weiss and Curie consists in supposing that the local magnetic field \mathbf{B}_{loc} felt by a magnetic moment is due to both the external magnetic field \mathbf{B} and the magnetic field generated by the ensemble of the magnetic moments, supposed to be proportional to the magnetization. That is

$$\mathbf{B}_{\text{loc}} = \mathbf{B} + \lambda \mathbf{M}$$



- Using the result of Ex. 9.4c., write the magnetization of an ensemble of spins under the presence of an external field $\mathbf{B} = B_0 \mathbf{u}_z$ by considering now that each spin feels the local field \mathbf{B}_{loc} .

- In the case of high temperatures, show that the magnetic susceptibility follows the Curie-Weiss law:

$$\chi_m = \frac{C}{T - T_C}$$

and determine the values of C and T_C .

- Going back to the general case, under what conditions can the magnetization \mathbf{M} be different from zero in the absence of an external field \mathbf{B} ?

Solution

- Assuming that the magnetization writes $\mathbf{M} = M \mathbf{u}_z$, a generalization of the magnetization found in Ex. 9.4c. gives

$$M = n \langle m \rangle = n \mu_B \tanh \left(\frac{\mu_B (B_0 + \lambda M)}{k_B T} \right).$$

- At high temperatures,

$$\tanh \left(\frac{\mu_B (B_0 + \lambda M)}{k_B T} \right) \approx \frac{\mu_B (B_0 + \lambda M)}{k_B T},$$

so we can write

$$M \approx \frac{n\mu_B^2(B_0 + \lambda M)}{k_B T} = \frac{n\mu_B^2 B_0}{k_B T} + \frac{n\mu_B^2 \lambda M}{k_B T}.$$

Solving for M and considering the limit of a weak susceptibility ($\chi_m \ll 1$)

$$M = \frac{\frac{n\mu_B^2 B_0}{k_B T}}{1 - \frac{n\mu_B^2 \lambda}{k_B T}} \approx \frac{B_0}{\mu_0} \chi_m,$$

we find the magnetic susceptibility

$$\chi_m = \frac{\frac{n\mu_0 \mu_B^2}{k_B T}}{1 - \frac{n\mu_B^2 \lambda}{k_B T}} = \frac{\frac{n\mu_0 \mu_B^2}{k_B}}{T - \frac{n\mu_B^2 \lambda}{k_B}}.$$

We identify the temperature $T_C = \frac{n\mu_B^2 \lambda}{k_B}$ and $C = n\mu_0 \mu_B^2 / k_B$ and we see that χ_m follows the Curie-Weiss law:

$$\chi_m = \frac{C}{T - T_C}.$$

Note that for $T \gg T_C$ this reduces to the result obtained in Ex. 9.4. As it will be shown below, T_C represents a critical temperature below which another magnetic phase emerges: ferromagnetism.

c. If $\mathbf{B} = \mathbf{0}$, the magnetization is given by

$$M = n\langle m \rangle = n\mu_B \tanh\left(\frac{\lambda\mu_B M}{k_B T}\right)$$

and this equation needs to be solved in a self-consistent way. We can rewrite it in terms of the critical temperature $T_C = \frac{n\mu_B^2 \lambda}{k_B}$ so that

$$\frac{M}{n\mu_B} = \tanh\left(\frac{M/(n\mu_B)}{T/T_C}\right)$$

and we see that, apart from the solution $M = 0$ (which corresponds to the paramagnetic phase), there exists a solution with $M \neq 0$ only if $T < T_C$. The material can be magnetized even in the absence of an external magnetic field. The temperature T_C is called Curie's temperature and marks the transition between paramagnetism and ferromagnetism.

9.9 Boundary conditions for the magnetic field

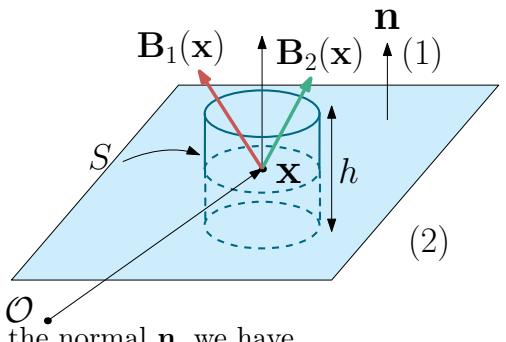
The fundamental laws of magnetostatics

$$\begin{aligned}\nabla \cdot \mathbf{B}(\mathbf{x}) &= 0, \\ \nabla \times \mathbf{H}(\mathbf{x}) &= \mathbf{J}(\mathbf{x}),\end{aligned}$$

provide information about the behavior of the fields at the interface that separates two magnetic media of different permeability.

9.9.1 The normal component of the magnetic field is continuous at an interface

Consider the situation shown in the figure below, where \mathbf{x} is a point at the interface between media 1 and 2, and let $\mathbf{B}_1(\mathbf{x})$ and $\mathbf{B}_2(\mathbf{x})$ be the magnetic field at a point infinitely close to \mathbf{x} in the region 1 and 2, respectively. We define \mathbf{n} , the normal to the surface that separates the two media, pointing from region 1 to region 2.



Taking a closed cylindrical surface S whose axis is parallel to the normal \mathbf{n} , we have

$$\nabla \cdot \mathbf{B} = 0 \rightarrow \iint_S \mathbf{B}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) dS(\mathbf{x}) = 0.$$

If the length h of the cylinder tends to zero, and if the area A of the cylinder caps is small enough so that we consider the magnetic field to be constant at all points of the caps, the flux integral essentially reduces to

$$S (\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n} = 0$$

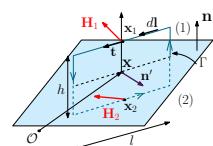
Consequently, at every point \mathbf{x} at the interface between two media:

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n}|_{\mathbf{x}} = 0. \quad (9.13)$$

That is, the **normal** component of the magnetic field is continuous across an interface.

9.9.2 The tangential component of the magnetic excitation is discontinuous at a current carrying surface

Now we consider a closed rectangular circuit Γ of length l and width h , oriented such that the normal to the planar surface $S(\Gamma)$ enclosed by Γ is $\mathbf{n}' = \mathbf{n} \times \mathbf{t}$, with \mathbf{t} the unit vector tangent to the interface along the horizontal segment of Γ contained in region 1.



Ampère's theorem integrated over the contour Γ reads

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{l} = \iint_{S(\Gamma)} \mathbf{J} \cdot (\underbrace{\mathbf{n} \times \mathbf{t}}_{\mathbf{n}'} dS.$$

If the width h of the loop tends to zero, then $\mathbf{x}_1 \rightarrow \mathbf{x}_2 = \mathbf{x}$, and only the horizontal segments of Γ contribute to the circulation of \mathbf{H} . If l is small enough so that the field is constant along the curve:

$$l \mathbf{t} \cdot (\mathbf{H}_1(\mathbf{x}) - \mathbf{H}_2(\mathbf{x})) = \lim_{h \rightarrow 0} \iint_{S(\Gamma)} \mathbf{J} \cdot \mathbf{n}' dS.$$

The term on the right hand side is zero when $h \rightarrow 0$ for any bulk current densities contained in regions 1 and 2. However, this term may differ from zero if there is a confined current density \mathbf{J}_S at the interface (in units of A m^{-1}), in which case

$$\lim_{h \rightarrow 0} \iint_{S(\Gamma)} \mathbf{J} \cdot \mathbf{n}' dS = l \mathbf{J}_S(\mathbf{x}) \cdot \mathbf{n}'$$

Finally,

$$\underbrace{\mathbf{t}}_{\mathbf{n}' \times \mathbf{n}} \cdot (\mathbf{H}_1(\mathbf{x}) - \mathbf{H}_2(\mathbf{x})) = \mathbf{J}_S(\mathbf{x}) \cdot \mathbf{n}'.$$

Using the identity $(\mathbf{n}' \times \mathbf{n}) \cdot (\mathbf{H}_1 - \mathbf{H}_2) = (\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2)) \cdot \mathbf{n}'$, we obtain

$$\{\mathbf{n} \times (\mathbf{H}_1(\mathbf{x}) - \mathbf{H}_2(\mathbf{x}))\} \cdot \mathbf{n}' = \mathbf{J}_S(\mathbf{x}) \cdot \mathbf{n}'.$$

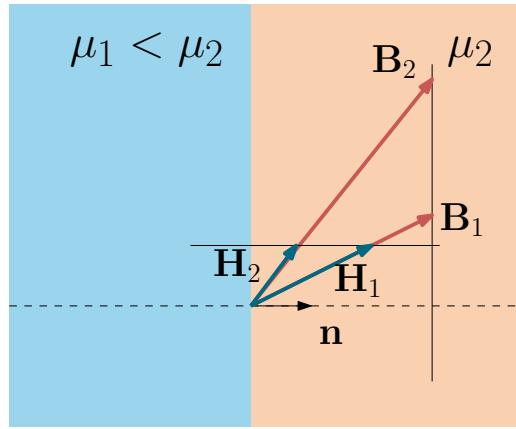
Since this is true for any \mathbf{n}' tangent to the interface, i.e., for any rotation of the contour Γ around the vertical axis \mathbf{n} , we conclude that for every \mathbf{x} at the interface between two media

$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2)|_{\mathbf{x}} = \mathbf{J}_S(\mathbf{x}), \quad (9.14)$$

where \mathbf{n} is the direction normal to the interface pointing from 2 to 1. In particular, in the absence of a current density at the interface, the tangential component of the magnetic excitation field \mathbf{H} is continuous.

A consequence of this continuity is that a high magnetic permeability allows the confinement of the magnetic field. Indeed, consider the case of an interface between air ($\mu_{\text{air}} = \mu_0$) and a material with high permeability, like for example a ferromagnet for which $\mu_{\text{Fe}} \gg \mu_0$. In the absence of surface currents, Eq. (9.14) states that the tangential component $\mathbf{H}_{||}$ of \mathbf{H} must be continuous at the interface, that is

$$\frac{\mathbf{B}_{||}^{\text{Fe}}}{\mu_{\text{Fe}}} = \frac{\mathbf{B}_{||}^{\text{air}}}{\mu_{\text{air}}} \quad \rightarrow \quad |\mathbf{B}_{||}^{\text{Fe}}| = \frac{\mu_{\text{Fe}}}{\mu_{\text{air}}} |\mathbf{B}_{||}^{\text{air}}| \gg |\mathbf{B}_{||}^{\text{air}}|$$



and so the tangential component of the magnetic field practically vanishes when passing from a ferromagnetic medium to a non-magnetic medium.

Suggested Exercises: Ex. 9.6 & 9.7 'Magnetic field of a magnetized sphere' and 'Magnetic shielding' in F. Cadiz's textbook

9.10 The magnetic free energy

We have established in chapter 8 the expression for the density of free energy, that is the amount of work necessary to establish a magnetic field \mathbf{B} , or equivalently, a vector potential and current distribution. We will generalize this expression to the case of magnetic media, that is, we are looking for the energy to be provided by an external operator to establish the magnetic induction \mathbf{H} and magnetic field \mathbf{B} in a magnetic medium that is initially not magnetized.

9.10.1 Differential expression for the volume density of free energy

The general expression for the vector potential in a magnetic medium

$$\mathbf{A}(\mathbf{x}) = \iiint_{\Omega} \frac{\mu_0}{4\pi} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \iiint_{\Omega} \frac{\mu_0}{4\pi} \frac{\mathbf{J}_M(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \iint_{\partial\Omega} \frac{\mu_0}{4\pi} \frac{\mathbf{J}_{S,M}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} dS,$$

shows that there is in general a nonlinear relation between \mathbf{A} and the conduction current density \mathbf{J} , due to the magnetization current density, \mathbf{J}_M in Ω and $\mathbf{J}_{S,M}$ in $\partial\Omega$.

Assume that we have an established steady state regime with conduction current density $\mathbf{J}(\mathbf{x}')$ for $\mathbf{x}' \in \Omega$ and corresponding vector potential $\mathbf{A}(\mathbf{x})$. If the conduction current varies by an infinitesimal amount to become $\mathbf{J}(\mathbf{x}') + \delta\mathbf{J}(\mathbf{x}')$, the vector potential will vary as

$$\mathbf{A}(\mathbf{x}) \rightarrow \mathbf{A}(\mathbf{x}) + \delta\mathbf{A}(\mathbf{x}),$$

at every point of space, and in particular at every point $\mathbf{x}' \in \Omega$. In result, an electromotive field will appear:

$$\mathbf{E}_{emf} = -\frac{\partial \mathbf{A}}{\partial t}.$$

The magnetic energy of the system corresponds to the work exerted by a virtual operator to withstand this electromotive field and maintain the volume Ω in a steady state. The operator must then exert an opposite field $-\mathbf{E}_{\text{emf}} = +\frac{\partial \mathbf{A}(\mathbf{x}')}{\partial t}$ in order to maintain the steady regime. The work of the operator to maintain the conduction currents is then equal to the work of the force $-dq\mathbf{E}_{\text{emf}}$ on an infinitesimal (free) charge $dq = \rho(\mathbf{x}')d^3x'$, which is obtained by taking the dot product of the force with the infinitesimal displacement vdt of the electrons, that is, per unit volume

$$\frac{\delta W}{d^3x'} = \rho(\mathbf{x}') \frac{\partial \mathbf{A}(\mathbf{x}')}{\partial t} \cdot \mathbf{v}(\mathbf{x}') dt = \rho(\mathbf{x}') \delta \mathbf{A}(\mathbf{x}') \cdot \mathbf{v}(\mathbf{x}') = \mathbf{j}(\mathbf{x}') \cdot \delta \mathbf{A}(\mathbf{x}')$$

This work represents the volume density of free energy: $df = \mathbf{J} \cdot \delta \mathbf{A}$. The magnetic free energy is then obtained by integration over the volume of the magnetic medium

$$dF = \iiint_{\Omega} \mathbf{j}(\mathbf{x}') \cdot \delta \mathbf{A}(\mathbf{x}') d^3x'$$

This integral can be extended to the entire space since there is no conduction current outside of Ω :

$$dF = \iiint_{\mathbb{R}^3} \mathbf{j}(\mathbf{x}') \cdot \delta \mathbf{A}(\mathbf{x}') d^3x'$$

Now we use a relation of vectorial calculus:

$$\nabla \cdot (\mathbf{H} \times \delta \mathbf{A}) = -\mathbf{H} \cdot \nabla \times \delta \mathbf{A} + \delta \mathbf{A} \cdot \nabla \times \mathbf{H},$$

and we identify $\nabla \times \delta \mathbf{A}$ from the definition of the magnetic field by means of the vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$, as well as $\nabla \times \mathbf{H}$ from Maxwell-Ampère's equation in steady regime:

$$\delta \mathbf{B} = \nabla \times \delta \mathbf{A}, \quad \text{and} \quad \nabla \times \mathbf{H} = \mathbf{J}.$$

From

$$\mathbf{J} \cdot \delta \mathbf{A} = \nabla \cdot (\mathbf{H} \times \delta \mathbf{A}) + \mathbf{H} \cdot \delta \mathbf{B},$$

we find the infinitesimal free energy

$$dF = \underbrace{\iiint_{\mathbb{R}^3} \nabla \cdot (\mathbf{H} \times \delta \mathbf{A}) d^3x'}_{\iint_{S \rightarrow \infty} (\mathbf{H} \times \delta \mathbf{A}) \cdot \mathbf{n} dS = 0} + \iiint_{\mathbb{R}^3} \mathbf{H} \cdot \delta \mathbf{B},$$

where the first integral on the right-hand side can be transformed into a surface integral by Ostrogradsky's theorem and vanishes as the product of the magnetic induction and the vector potential decays faster than $1/r^3$. Finally

$$dF = \iiint_{\mathbb{R}^3} \mathbf{H} \cdot \delta \mathbf{B},$$

leading to the expression for the differential of the volume density of free energy $f(\mathbf{B})$:

$$df = \mathbf{H} \cdot d\mathbf{B}.$$

H and **B** are conjugated variables. The free energy is a state function, which can be expressed as a function of only one of the conjugated variables, that is, **B**. To do this, we need an equation of state for the magnetic media. Before giving an example, we express general properties for the magnetic free energy.

9.10.2 Properties for the magnetic free energy

From the expression for the magnetic induction, we have

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}),$$

hence by differentiation,

$$d\mathbf{B} = \mu_0(d\mathbf{H} + d\mathbf{M}),$$

and the differential for the free energy density exhibits two terms:

$$df = \underbrace{\mu_0 \mathbf{H} \cdot d\mathbf{H}}_{\text{energy to establish } \mathbf{H}} + \underbrace{\mu_0 \mathbf{H} \cdot d\mathbf{M}}_{\text{energy to establish magnetization}}.$$

These two terms correspond to the energy to establish the magnetic induction, and the energy to establish the magnetization in the medium, respectively. If \mathbf{M} can be neglected with respect to \mathbf{H} , we have $\mathbf{B} \simeq \mu_0 \mathbf{H}$ and

$$df = \mathbf{B} \cdot d\mathbf{M}.$$

Only the second term to establish the magnetization is relevant in the differential for the free energy density.

9.10.3 Free energy vs free enthalpy

Having identified the conjugated variables \mathbf{B} and \mathbf{H} from the expression for the free energy density $df = \mathbf{H} \cdot d\mathbf{B}$, we can define the free enthalpy by performing the Legendre transform

$$g = f - \mathbf{H} \cdot \mathbf{B},$$

whose differential is

$$dg = -\mathbf{B} \cdot d\mathbf{H}.$$

In other words, the free enthalpy $g(\mathbf{H})$ is expressed as a natural function of \mathbf{H} . A constitutive relation, i.e., an equation of state (e.g. $\mathbf{H}(\mathbf{B})$ or $\mathbf{H}(\mathbf{M})$), is required to find expressions for the magnetic free energy or enthalpy.

9.10.4 Free energy for linear, homogeneous, isotropic media

In the case of paramagnetic or diamagnetic materials, \mathbf{B} and \mathbf{H} are linked by $\mathbf{B}(\mathbf{x}) = \mu \mathbf{H}(\mathbf{x})$ where the permeability of the medium $\mu = \mu_0 \mu_r$ is the product of the permeability of vacuum by μ_r , the relative magnetic permeability. We can then find the equation of state $\mathbf{M}(\mathbf{H})$ in the following way:

$$\left\{ \begin{array}{l} \mathbf{B} = \mu \mathbf{H} \\ \frac{\mathbf{B}}{\mu_0} = \mathbf{H} + \mathbf{M} \end{array} \right. \Rightarrow \left(\frac{\mu}{\mu_0} - 1 \right) \mathbf{H} = \mathbf{M},$$

$$\mu_r = 1 + \chi$$

$$\mathbf{M} = \chi \mathbf{H}$$

where we recognize the magnetic susceptibility χ of the medium. Keeping in mind that for paramagnetic and diamagnetic media, the susceptibility is small

$$\left. \begin{array}{ll} \text{Paramagnetism:} & \chi > 0 \quad \chi \sim 10^{-3} \\ \text{Diamagnetism:} & \chi < 0 \quad \chi \sim -10^{-5} \end{array} \right\} |\chi| \ll 1 \rightarrow \text{No local field,}$$

we find an equation of state linking the magnetization to the magnetic field

$$\boxed{\mathbf{M} = \chi \mathbf{H} \quad \Rightarrow \quad \mathbf{M} = \chi \frac{\mathbf{B}}{\mu_0}.}$$

This shows that \mathbf{M} , \mathbf{B} and \mathbf{H} are parallel. In addition, the sign of the susceptibility tells us that the magnetization points in the same direction as the magnetic field for paramagnetic materials, and in the opposite direction for diamagnetic materials.

The volume density of free energy can then be integrated easily. From $df = \mathbf{H} \cdot d\mathbf{B}$ with $\mathbf{B} = \mu \mathbf{H}$ (linear medium), we find

$$\boxed{f = \frac{1}{2} \mathbf{H} \cdot \mathbf{B} \quad \text{or} \quad f(\mathbf{B}) = \frac{\mathbf{B}^2}{2\mu}.}$$

For a magnetized domain, the magnetic free energy is then

$$\boxed{F = \iiint_{\mathbb{R}^3} \frac{\mathbf{B}^2(\mathbf{x}')}{2\mu} d^3x'}$$

9.11 Summary and essential formulas

- At the atomic scale, the orbital motion of electrons in atoms leads to an orbital magnetic moment $\mathbf{m}_L = -\frac{\mu_B}{\hbar}\mathbf{L}$, where \mathbf{L} is the electron angular momentum, $\mu_B \approx 9 \times 10^{-24} \text{ A m}^2$ is Bohr's magneton and \hbar is the reduced Planck constant. In addition, every electron has a spin, that is an intrinsic angular momentum \mathbf{S} with no classical analogue, leading to a spin magnetic moment $\mathbf{m}_S \approx -2\frac{\mu_B}{\hbar}\mathbf{S}$.
- The magnetic moment of an atom results from the contributions of orbital and spin magnetic moments of its constituent particles. Under the presence of an external magnetic field \mathbf{B} , a magnetic moment precesses in the plane perpendicular to \mathbf{B} at a characteristic frequency, called the Larmor frequency. At the same time, it rotates to eventually align with the magnetic field after a typical time T_1 called the spin-lattice relaxation time. It corresponds to the time required for the spin to reach thermal equilibrium in the presence of \mathbf{B} . A magnetic field is thus capable of inducing a magnetization in matter.
- Magnetic materials can be classified into three main categories depending on their response to an external magnetic field:
 - Diamagnetics: these corresponds to insulators for which the electronic shells are full, so that the atoms do not have a net magnetic moment and their magnetic activity is due to a modification of the atomic orbital in the presence of a magnetic field. They are repelled by an external field.
 - Paramagnetics: they are media for which the magnetic moment of either free electrons or atoms with unsaturated electronic shells align along the direction of an external magnetic field to reach the most energetically stable configuration. They become magnetized and are attracted by an external magnetic field, but whenever the latter is turned off, they loose their magnetization due to thermal disorder.
 - Ferromagnetics: they are media that exhibit a remanent magnetization even after the external magnetic field is turned off. This is due to a strong interaction between neighboring magnetic moments, and has a pure quantum origin: because of the Pauli principle, the Coulomb repulsion between two electrons depends on the relative orientation of their spins.
- A magnetic material is characterized by the magnetization vector \mathbf{M} , which represents the average density of magnetic dipole moment, i.e., the averaged magnetic dipole moment per unit volume.

$$\mathbf{M} = \frac{d\mathbf{m}}{d^3x}$$

- A magnetic medium Ω of magnetization \mathbf{M} generates a magnetic field that is equivalent to that of a volume having a stationary and localized current, composed of a current density \mathbf{J}_V within

its volume and \mathbf{J}_S at its boundary $\partial\Omega$. They are given by

$$\begin{aligned}\mathbf{J}_M(\mathbf{x}) &= \nabla \times \mathbf{M}(\mathbf{x}), \quad \text{for } \mathbf{x} \in \Omega, \\ \mathbf{J}_{S,M}(\mathbf{x}) &= \mathbf{M}(\mathbf{x}) \times \mathbf{n}(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial\Omega,\end{aligned}$$

where \mathbf{n} is normal to $\partial\Omega$.

- Ampère's law in a material medium is written:

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{l} = I_{\text{free}} \Leftrightarrow \nabla \times \mathbf{H} = \mathbf{J}_{\text{free}},$$

where I_{free} denotes the total free (conduction) current enclosed by Γ and \mathbf{H} is the magnetic excitation vector, defined as

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}.$$

- In a homogeneous and isotropic paramagnetic or diamagnetic medium, the magnetization is proportional to \mathbf{H} , that is $\mathbf{M} = \chi_m \mathbf{H}$, with $\chi_m \sim -10^{-5} < 0$ for a diamagnetic medium, and $\chi_m \sim 10^{-3} > 0$ for a paramagnetic medium. In a ferromagnetic medium, however, the magnetic susceptibility $\chi_m(\mathbf{H})$ is given by a multivalued function of \mathbf{H} that depends on the history and may give rise to the phenomenon of hysteresis.
- The \mathbf{H} vector can be written:

$$\mathbf{H} = \frac{\mathbf{B}}{\mu},$$

where $\mu = \mu_0(1 + \chi_m)$ is the permeability of the medium. Everything happens as in vacuum, provided that the permeability is renormalized, which may be less than or greater than μ_0 . The fundamental equations of electrostatics in matter are then written:

$$\nabla \times \mathbf{B} = \mu \mathbf{J}_{\text{free}}, \quad \nabla \cdot \mathbf{B} = 0,$$

where \mathbf{J}_{free} represents the density of free (conduction) current.

- At the interface between two media:
 - The normal component of the magnetic field is continuous

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n}|_{\mathbf{x}} = 0.$$

where \mathbf{x} denote any point at the interface and \mathbf{n} is the normal to the interface pointing from medium (2) to medium (1).

- The tangential component of the magnetic induction \mathbf{H} is discontinuous

$$\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) |_{\mathbf{x}} = \mathbf{J}_S(\mathbf{x}),$$

where \mathbf{J}_S denotes the surface current density at the interface and \mathbf{n} is the normal to the interface pointing from medium (2) to medium (1).

Chapter 10

Maxwell's equations and wave propagation in matter

Introduction

We begin this chapter by generalizing Ampère's theorem in order to make it compatible with the charge conservation law in the time-dependent regime. This generalized law, known as Ampère-Maxwell's law, together with Gauss's law, Faraday's law and Thomson's law constitute a complete set of coupled equations between the electric and magnetic fields, known as Maxwell's equations in honor to James Clerk Maxwell who put them together¹ in 1864. Maxwell's equation not only synthesize all the known electrostatic and magnetostatic phenomena discussed so far, but also establishes a profound relationship between them. Indeed, they predict the existence of electric and magnetic disturbances which propagate together, called *electromagnetic waves* propagating at the speed $c = 1/\sqrt{\epsilon_0\mu_0} = 3 \times 10^8$ m/s, close to the value for the speed of light measured by Fizeau and Foucault, so that Maxwell proposed that light was an electromagnetic phenomenon.

The prediction of such waves was confirmed experimentally between 1886 and 1889 by Heinrich Hertz who demonstrated the existence of what is now called radio waves that propagate at the speed of light. This was a profound leap in our understanding of nature: radio waves and light were both a form of electromagnetic radiation obeying Maxwell's equations.

From Maxwell's equations we will derive the propagation laws of such waves and discuss the mathematical structure (transversality, polarization) of the solutions, firstly in vacuum, and in a second part in matter (dielectrics and conductors). Special attention will be paid to the transport of electromagnetic energy which follows a conservation equation known as Poynting's theorem. While the present chapter mainly investigates the propagation of electromagnetic waves in ideal cases, that is, plane waves and sinusoidal plane waves propagate in matter without absorption, the next chapter will treat wave propagation in conductors and dielectrics where absorption due to the inertia of charged particles is no longer neglected.

¹In reality Maxwell wrote down 20 equations that were later reduced to only four by Oliver Heaviside, after the development of the vectorial calculus.

10.1 Ampère's theorem and displacement current

So far, we have established the following laws that summarize the electrostatic, magnetostatic, and Faraday's induction phenomena:

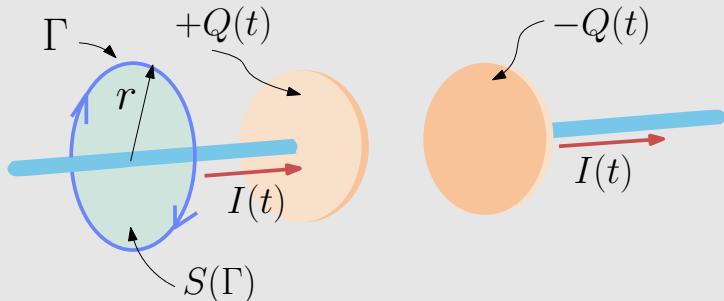
$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$	Gauss's law,
$\nabla \cdot \mathbf{B} = 0$	Thomson's law,
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday's law,
$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$	Ampère's law.

(10.1)

This set of equations is, however, incomplete. The problem comes from the fact that Ampère's law is inconsistent in the presence of time-dependent current densities. In order to illustrate this, let us consider the example below.

Example 10.1 - Problem with Ampère's law

The following is a problematic example when applying Ampère's law in a situation where the current is not stationary. Consider a parallel plate capacitor that is being charged with a current $I(t) = dQ(t)/dt$. The goal is to determine the magnetic field at a distance r from the cable, sufficiently far from the plates and close to the cable so we can assume a cylindrical symmetry of the current density. A possible way to calculate the magnetic field is to use Ampère's law integrated over a circular path Γ of radius r as shown in the figure below.



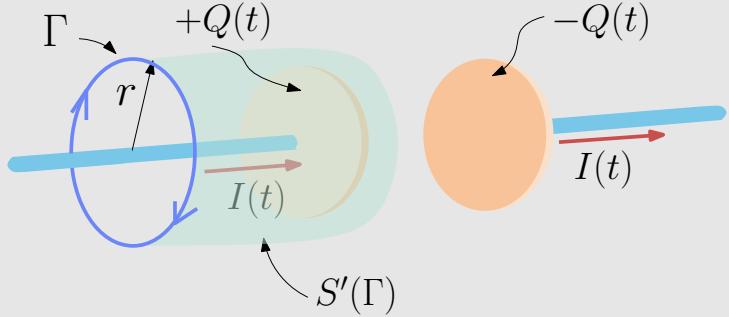
We obtain

$$\oint_{\Gamma} \mathbf{B} \cdot d\mathbf{x} = 2\pi r B(r) = \mu_0 \Phi_{S(\Gamma), \mathbf{J}} = \mu_0 \iint_{S(\Gamma)} d\mathbf{S}(\mathbf{x}) \cdot \mathbf{J}(\mathbf{x}),$$

where $S(\Gamma)$ is any surface whose bounding curve is Γ . For example, choosing the flat surface coplanar and enclosed by Γ , the flux of \mathbf{J} through $S(\Gamma)$ corresponds to the current I and we obtain the well known result from magnetostatics

$$2\pi r B(r) = \mu_0 I.$$

However, we could have chosen instead the surface $S'(\Gamma)$ as shown in the following figure.



In this case, there is no flux of \mathbf{J} through $S'(\Gamma)$ and a contradiction is obtained:

$$2\pi r B(r) = 0.$$

The problem arises from the fact that Ampère's law is not compatible with time-varying current densities.

Ampère's law is inconsistent with the conservation law for the electric charge:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

Indeed, taking the divergence of Ampère's law in its differential form gives

$$\nabla \cdot (\nabla \times \mathbf{B}) = \mu_0 \nabla \cdot \mathbf{J} = 0,$$

since the divergence of a rotational is always zero, and so $\nabla \cdot \mathbf{J} = 0$. While this is not a problem in the stationary regime (ρ independent of time), it is in contradiction with the conservation of charge in the time-dependent regime, for which

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \neq 0.$$

This problem was solved by Maxwell, who added to Ampère's law a second current density term \mathbf{J}_D , called *displacement current* for reasons that will become clear in the section about Maxwell's equations in matter. This current is such that the total current $\mathbf{J}_T = \mathbf{J} + \mathbf{J}_D$, and not \mathbf{J} , has zero divergence.

10.1.1 Displacement current

Maxwell corrected Ampère's law by adding a term so as to make the law consistent with charge conservation:

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \mathbf{J}_D), \quad (10.2)$$

where the *displacement current* \mathbf{J}_D can be inferred by using both charge conservation and Gauss's law. By taking the divergence on both sides of Eq. (10.2), we now obtain

$$\nabla \cdot (\nabla \times \mathbf{B}) = \mu_0(\nabla \cdot \mathbf{J} + \nabla \cdot \mathbf{J}_D) = 0.$$

Since the charge continuity equation imposes $\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}$, then,

$$\nabla \cdot \mathbf{J}_D = \frac{\partial \rho}{\partial t}.$$

Taking the partial derivative with respect to time in Gauss's law, we find $\nabla \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{\epsilon_0} \frac{\partial \rho}{\partial t}$, and then, we see that it is possible to choose

$$\mathbf{J}_D = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (10.3)$$

It is this choice that was done by Maxwell in 1862 to cure the problem of Ampère's law in the time dependent regime and obtain the fourth Maxwell equation. This was a truly brilliant intuition since Maxwell did not know the existence of the electron and did not think in terms of charge conservation when he introduced the displacement current as *one of the chief peculiarities* of his theory. He postulated the set of fundamental equations unifying electric and magnetic phenomena; Maxwell's equations led to the prediction of electromagnetic waves propagating at the speed of light, which, according to A. Zangwill, might be one of the arguments that convinced him of the correctness of the displacement current. It is only after 1888 when Hertz discovered electromagnetic waves behaving as predicted by Maxwell's theory that the rest of the world accepted Maxwell's postulate as correct, that is, the fact that Maxwell's equations are the fundamental laws describing electric and magnetic phenomena in a unified framework and that light is an electromagnetic wave.

Definition 10.1: Ampère-Maxwell's law

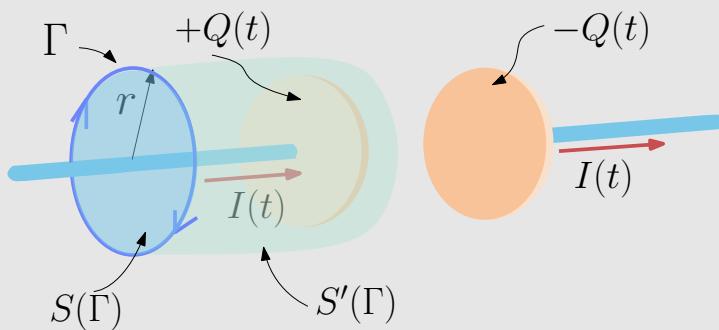
Ampère-Maxwell's law reads

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \quad (10.4)$$

In the case of fields that do not depend on time, this law reduces to Ampère's law. The introduction of the displacement current $\mathbf{J}_D = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$ in the second term represents a major contribution from Maxwell to the theory of electromagnetism, since it allowed him to predict the existence of electromagnetic waves.

Example 10.2 - Charging capacitor and Ampère-Maxwell's law

Let us consider again Ex. 10.1. Assuming for simplicity that the separation between the plates is small enough compared to their lateral dimensions, the electric field \mathbf{E} (and therefore the displacement current \mathbf{J}_D) is totally confined between the plates. According to Ampère-Maxwell's law, the circulation of the magnetic field around the closed curve Γ equals the flux of $\mu_0(\mathbf{J} + \mathbf{J}_D) = \mu_0 \mathbf{J}_T$ over any^a surface whose bounding curve is Γ .



If we choose the planar surface $S(\Gamma)$, the flux is only due to \mathbf{J} and given by $\mu_0 \Phi_{S(\Gamma), \mathbf{J}} = \mu_0 I$. If instead we choose the surface $S'(\Gamma)$, the flux is due to \mathbf{J}_D and we have

$$\mu_0 \Phi_{S'(\Gamma), \mathbf{J}_D} = \mu_0 \epsilon_0 \frac{\partial}{\partial t} \Phi_{S'(\Gamma), \mathbf{E}}.$$

The flux of \mathbf{E} through $S'(\Gamma)$ is equal to the flux over the closed surface $S'(\Gamma) - S(\Gamma)$,^a which by Gauss's law equals $Q(t)/\epsilon_0$. Thus,

$$\mu_0 \Phi_{S(\Gamma), \mathbf{J}_D} = \mu_0 \frac{\partial Q(t)}{\partial t} = \mu_0 I,$$

so that Ampère-Maxwell's law gives the same result regardless of the surface that is chosen to calculate the flux of $\mathbf{J} + \mathbf{J}_D$:

$$\oint_{\Gamma} \mathbf{B} \cdot d\mathbf{x} = \mu_0 \Phi_{S(\Gamma), \mathbf{J}_T} = \mu_0 \Phi_{S'(\Gamma), \mathbf{J}_T} = \mu_0 I.$$

^aThe flux can be taken through *any* surface bounded by Γ because $\nabla \cdot \mathbf{J}_T = 0$.

^bThis is the union of $S'(\Gamma)$ and $-S(\Gamma)$, i.e., $S(\Gamma)$ with the opposite orientation to ensure that the normal to the total surface always points outward.

Suggested additional lectures: Ex.11.3 & 11.4 in F. Cadiz's textbook.

Definition 10.2: Maxwell's equations

Maxwell's equations are the following set of coupled equations relating the electric and magnetic fields (\mathbf{E} and \mathbf{B}) to their sources (charge density ρ and current density \mathbf{J}):

$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$	Gauss-Maxwell,
$\nabla \cdot \mathbf{B} = 0$	Thomson-Maxwell,
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday-Maxwell,
$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$	Ampère-Maxwell.

(10.5)

To the Maxwell equations, we must add the Lorentz force that is exerted on a particle of charge q of

velocity \mathbf{v} in an electromagnetic field,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

This set of laws, together with Newton's equation for the velocity of charged particles, gives us a complete description of classical electromagnetism. Before discussing the possible solutions to Maxwell's equations, a reminder of the wave equation and its general solutions is presented below.

10.2 D'Alembert's wave equation

The wave equation in three dimensions (3D) corresponds to the following partial differential equation for a scalar field $F(\mathbf{x}, t)$,

$$\nabla^2 F(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 F(\mathbf{x}, t)}{\partial t^2} = 0,$$

where $\mathbf{x} \equiv (x, y, z)$ and $\nabla^2 F \equiv \partial^2 F / \partial x^2 + \partial^2 F / \partial y^2 + \partial^2 F / \partial z^2$. For simplicity, let us first consider the one-dimensional (1D) case of a wave propagating in the x -direction,

$$\frac{\partial^2 F(x, t)}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 F(x, t)}{\partial t^2} = 0. \quad (10.6)$$

As demonstrated by d'Alembert in 1747, any function of the form $F(x \pm ct)$, where $X \rightarrow F(X)$ is any twice differentiable function of a single variable, is solution to the wave equation (10.6).

Proof

Indeed, consider the change of variable $X = x \pm ct$. Partial derivatives of X with respect to x or t read $\frac{\partial X}{\partial x} = 1$ and $\frac{\partial X}{\partial t} = \pm c$. Using the chain rule, we find

$$\frac{\partial F(X)}{\partial x} = \frac{\partial F(X)}{\partial X} \frac{\partial X}{\partial x} = \frac{\partial F(X)}{\partial X},$$

and so,

$$\frac{\partial^2 F(X)}{\partial x^2} = \frac{\partial^2 F(X)}{\partial X^2}.$$

Similarly, the partial derivative with respect to t reads

$$\frac{\partial F(X)}{\partial t} = \frac{\partial F(X)}{\partial X} \frac{\partial X}{\partial t} = \pm c \frac{\partial F(X)}{\partial X}$$

and so,

$$\frac{\partial^2 F(X)}{\partial t^2} = \frac{\partial}{\partial t} \left(\pm c \frac{\partial F(X)}{\partial X} \right) = \pm c \frac{\partial^2 F(X)}{\partial X^2} \frac{\partial X}{\partial t} = c^2 \frac{\partial^2 F(X)}{\partial X^2}.$$

We see that $F(X) = F(x \pm ct)$ is solution to d'Alembert's wave equation,

$$\frac{\partial^2 F(X)}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 F(X)}{\partial t^2}.$$

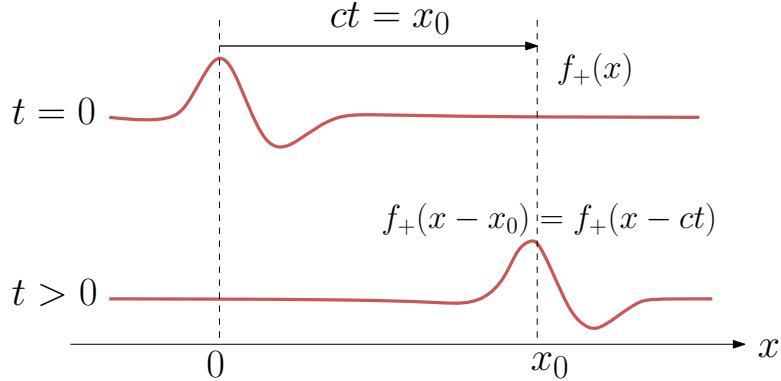
The latter equation being linear, the superposition principle applies and we conclude that a possible solution is given by

$$F(x, t) = Af_+(x - ct) + Bg_-(x + ct), \quad (10.7)$$

where f_+ and g_- are twice differentiable functions of a single variable, and where the constants A and B depend on boundary and initial conditions.

• PHYSICAL INTERPRETATION

Consider a solution of the form $f_+(x - ct)$ with $c > 0$, which at $t = 0$ represents a disturbance $f_+(x)$. After a time $t > 0$, the wave corresponds to $f_+(x - x_0)$, which is simply f_+ shifted towards the positive values of x by a quantity $x_0 = ct > 0$. We conclude that $f_+(x - ct)$ represents a disturbance that propagates without deformation towards positive values of x at a speed c . It is therefore called a progressive wave.



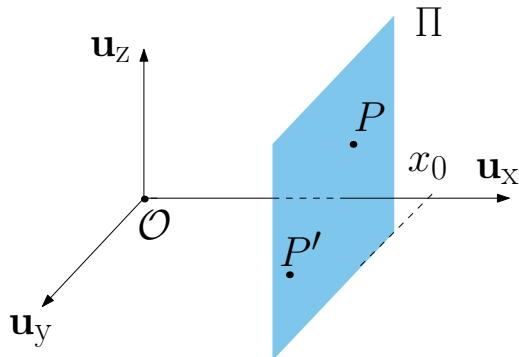
In the same way, we see that $g_-(x + ct)$ corresponds to a regressive wave, that is, a disturbance that propagates without deformation towards the direction of decreasing values of x . Note finally that the general solution $Af_+(x - ct) + Bg_-(x + ct)$, which is always a superposition of progressive and regressive waves, is not necessarily a traveling wave. It could be, for example, a standing wave.

10.2.1 Plane waves and wave fronts

Consider again the wave equation in three dimensions,

$$\nabla^2 F(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 F(\mathbf{x}, t)}{\partial t^2} = 0.$$

The wave F is said to be a plane wave propagating along the direction of propagation \mathbf{n} if, for every t , the function F has the same value at every point of a plane Π perpendicular to \mathbf{n} . For example, for a plane wave propagating in the x -direction, this means that F does not depend on y or z , but only on x . This is represented in the following figure; F has the same value at P and at P' , both belonging to a plane $x = x_0$.



In this case, the definition of a plane wave in three dimensions results in F satisfying the 1D d'Alembert equation

$$\frac{\partial^2 F}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 F}{\partial t^2} = 0.$$

10.2.2 Sinusoidal plane waves and wavefronts

Consider a plane wave that propagates along the x -axis. A preferred solution to d'Alembert's equation corresponds to sinusoidal plane waves. A progressive sinusoidal wave reads

$$f_+(x, t) = \underbrace{A_1}_{\text{amplitude}} \cos(\underbrace{kx - \omega t + \phi_1}_{\text{phase}})$$

and similarly, a regressive sinusoidal wave reads

$$g_-(x, t) = A_2 \cos(kx + \omega t + \phi_2)$$

At a given position $x = x_0$, the wave oscillates in time with a period

$$T = \frac{2\pi}{\omega},$$

whereas at a fixed time $t = t_0$, the wave oscillates in space with a spatial period (also called wavelength)

$$\lambda = \frac{2\pi}{k}.$$

We see then that ω and k correspond to the temporal and spatial frequencies of the wave, respectively. These quantities are not independent. Indeed, injecting f_+ and g_- into d'Alembert's equation leads to the following equation linking ω and k , also called the dispersion relation,

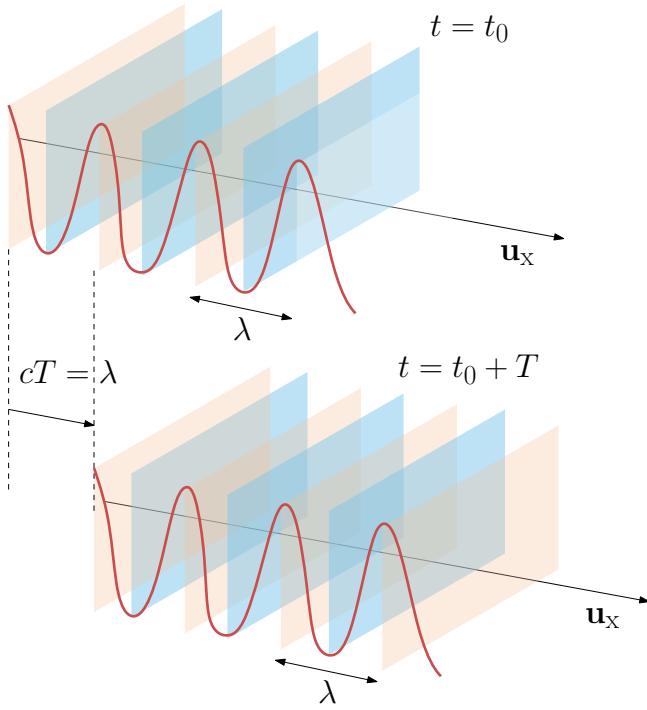
$\omega = kc.$

(10.8)

A wavefront is, at a given instant, a surface of constant phase. For a sinusoidal plane wave propagating along the x -axis, this means

$$kx + \phi = \text{constant} \rightarrow x = \text{constant}.$$

The following figure represents a progressive sinusoidal wave traveling along the x -direction, together with some wavefronts spaced by $\lambda/2$. According to the dispersion relation (10.8), during one temporal period T , the wave has traveled a distance $cT = \lambda$.



10.2.3 Spherical waves

Consider a wave with spherical symmetry. This would correspond to the case of a wave isotropically emitted from (or focused at) the origin of coordinates. For such a wave, which only depends on r , that is $F = F(r, t)$, the wave equation reads

$$\nabla^2 F(r, t) - \frac{1}{c^2} \frac{\partial^2 F(r, t)}{\partial t^2} = \frac{1}{r} \frac{\partial^2}{\partial r^2} (rF) - \frac{1}{c^2} \frac{\partial^2 F}{\partial t^2} = 0.$$

Writing $\Psi(r, t) = rF(r, t)$, the equation for Ψ reads

$$\frac{\partial^2 \Psi(r, t)}{\partial r^2} - \frac{1}{c^2} \frac{\partial^2 \Psi(r, t)}{\partial t^2}.$$

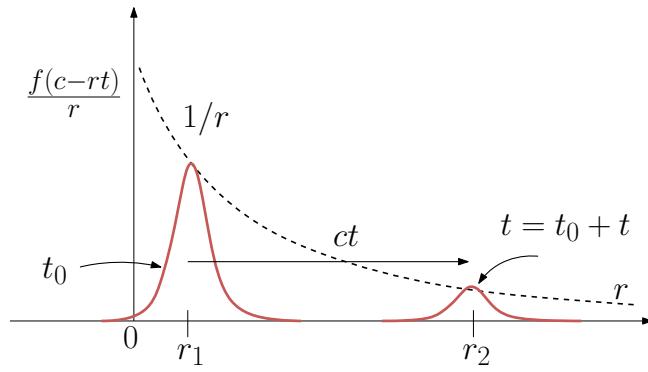
which is d'Alembert's one-dimensional wave equation, whose solution is in general a combination of progressive and regressive waves:

$$\Psi(r, t) = Af(r - ct) + Bg(r + ct).$$

Finally, going back to $F(r, t) = \Psi(r, t)/r$, the spherical wave is

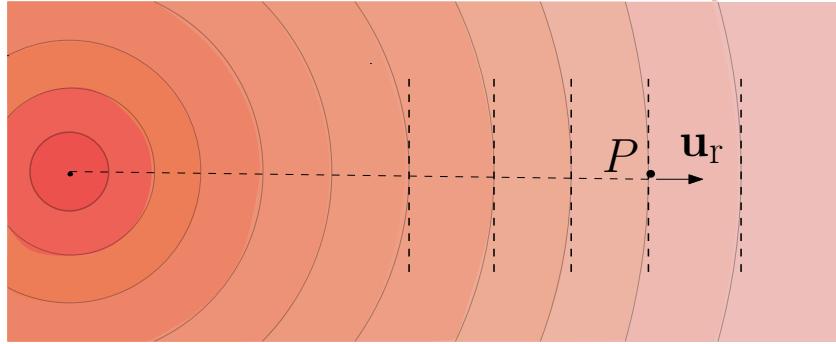
$$F(r, t) = \frac{A}{r} f(r - ct) + \frac{B}{r} g(r + ct).$$

Unlike plane waves, the amplitude of spherical waves changes during propagation. In electromagnetism, this reflects the conservation of energy as the wave propagates. The following figure shows the case of a progressive (along the radial direction) spherical wave of the form $f(r - ct)/r$. The shape of the wave is preserved, although its amplitude decreases as $1/r$ as the wave travels. The wave looks as if it were emitted from a point source at the origin.



Note that, at a given time, the amplitude of the wave takes the same value at all points such that $r = r_0$, that is, the wave is constant on any sphere centered at $r = 0$. Finally, a solution of the form $g(r+ct)/r$ seems to be emitted from a spherical surface far from the origin and to be focused at $r = 0$. The amplitude of such a wave then increases as r decreases.

If we choose a sinusoidal spherical wave $f(r-ct) = A \cos(kr - \omega t)$ we have $F(r, t) = \frac{A}{r} \cos\{k(r-ct)\}$ so that, at any given time, the wavefronts are spheres centered at the origin. At a point P sufficiently far from $r = 0$, these wavefronts can be locally approximated by planes perpendicular to the radial direction \mathbf{u}_r so that the wave behaves, locally, like a plane wave propagating along \mathbf{u}_r .



Sinusoidal plane waves can thus be used to describe the local behavior of spherical waves. Near $r = r_0$,

$$F(r_0 + dr, t) = \frac{A \cos(k(r_0 + dr) - \omega t)}{r_0 + dr} \approx \frac{A \cos(kdr - \omega t + kr_0)}{r_0} = A' \cos(kdr - \omega t + \phi).$$

10.2.4 Complex representation of sinusoidal waves

A sinusoidal wave of the form $u(x, t) = A \cos(kx - \omega t + \phi)$ may be written in a complex representation, thanks to Euler's identity ($e^{ix} = \cos x + i \sin x$), as the real part of the following complex wave,

$$\underline{u}(x, t) = \underline{A} e^{i(kx - \omega t)},$$

with $\underline{A} = Ae^{i\phi}$. Indeed,

$$u(x, t) = \operatorname{Re}\{\underline{u}(x, t)\} = \operatorname{Re}\{Ae^{i(kx - \omega t + \phi)}\} = A \cos(kx - \omega t + \phi).$$

The complex representation significantly simplifies calculations involving sinusoidal waves, since calculus with complex exponentials is easier compared to calculus with trigonometric formulas. The physical result of any linear operation $L(u)$ acting on u can then be obtained as $\operatorname{Re}\{L(\underline{u})\}$.

10.2.5 Why focus on sinusoidal waves?

A sinusoidal wave has an infinite extension in both space and time, and as such, it cannot represent on its own a realistic physical phenomenon. For instance, its finite and constant amplitude, in conjunction with its infinite extension, means that it carries an infinite energy. In addition, it may seem arbitrary to focus on sinusoidal plane waves since any twice differentiable function can be a solution of d'Alembert's wave equation. But in fact, the restriction to sinusoidal waves is fully justified by the following important result of Fourier analysis: any physical plane wave $f(x, t)$ (bounded in both space and time) can be written as a linear, (infinite) superposition of sinusoidal waves

$$f(x, t) = \int_{\mathbb{R}} F(k, t) e^{i(kx - \omega(k)t)} dk.$$

Sinusoidal plane waves are thus the building blocks of real waves, and due to the linearity of Maxwell's equations, the behavior of electromagnetic plane waves is enough to characterize the most general case.

10.3 Electromagnetic waves in vacuum

An immediate consequence of Maxwell's equations is the prediction of the existence of electromagnetic waves that propagate in vacuum at the speed of light $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 3 \times 10^8 \text{ m s}^{-1}$. This is possible because, in the absence of sources ($\rho = 0$ in Gauss-Maxwell's equation and $\mathbf{J} = \mathbf{0}$ in Ampère-Maxwell's equation), the electric and magnetic fields are coupled according to the two equations involving a curl,

$$\underbrace{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}}_{\text{time-varying } \mathbf{B} \text{ generates } \mathbf{E}}, \quad \underbrace{\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}}_{\text{time-varying } \mathbf{E} \text{ generates } \mathbf{B}}$$

which leads to the propagation of the fields. This prediction was confirmed experimentally by Hertz in 1888.

10.3.1 Propagation equation for the electric field

The propagation equation is obtained using the following identity, where \mathbf{F} is a vector field:

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla (\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}.$$

For the electric field, the double curl reads

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right) = -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}),$$

where we have used Maxwell-Faraday's law. Now using Ampère-Maxwell's law to replace $\nabla \times \mathbf{B}$, we obtain

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \left(\mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right).$$

Thus,

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = - \left(\mu_0 \frac{\partial}{\partial t} \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \right).$$

Finally, using Gauss's law $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, we obtain

$$\nabla^2 \mathbf{E} - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \mathbf{E} = \frac{\nabla \rho}{\epsilon_0} + \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (10.9)$$

Equation (10.9) is the propagation equation for the electric field, establishing a relationship between the spatial and time derivatives of the electric field, the charge density and the variation of the current density. Note that the coupling between the electric and magnetic fields is no longer explicit in this equation.

10.3.2 Propagation equation for the magnetic field

Taking the curl of Ampère-Maxwell's law yields

$$\nabla \times (\nabla \times \mathbf{B}) = \mu_0 \nabla \times \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial}{\partial t} \nabla \times \mathbf{E} = \nabla (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}.$$

Since $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E}$ can be replaced by means of Maxwell-Faraday's law, $\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{B}$, we obtain

$$\nabla^2 \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \mathbf{B} = -\mu_0 \nabla \times \mathbf{J}, \quad (10.10)$$

which corresponds to the propagation equation for the magnetic field.

10.3.3 Propagation in vacuum

Next, we consider the behavior of the fields away from the sources that generate them, i.e., in regions where $\rho = 0$ and $\mathbf{J} = \mathbf{0}$. We see then from Eqs. (10.9) and (10.10) that each component of both \mathbf{E} and \mathbf{B} satisfies the same d'Alembert wave equation,

$$\nabla^2 \mathbf{E} - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \mathbf{E} = 0, \quad (10.11)$$

$$\nabla^2 \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \mathbf{B} = 0, \quad (10.12)$$

where $\mu_0 \epsilon_0$ is the inverse of a squared speed, which corresponds to the speed of light

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \approx 3 \times 10^8 \text{ m s}^{-1}$$

10.3.4 Structure of electromagnetic plane waves

- TRANSVERSALITY

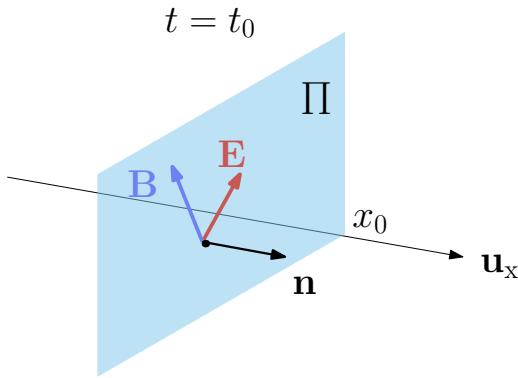
Consider an electromagnetic plane wave that propagates in vacuum along the x -axis, toward the region $x > 0$. This means that the fields are of the form $\mathbf{E}(x, t) = \mathbf{E}(x - ct)$ and $\mathbf{B}(x, t) = \mathbf{B}(x - ct)$. Let us define the change of variable $X = x - ct$. Gauss's law in vacuum implies

$$\nabla \cdot \mathbf{E} = \nabla X \cdot \frac{\partial \mathbf{E}}{\partial X} = 0$$

and since $X = x - ct$, $\nabla X = \mathbf{u}_x$ and so,

$$\nabla \cdot \mathbf{E} = \mathbf{u}_x \cdot \frac{\partial \mathbf{E}}{\partial X} = \frac{\partial E_x}{\partial X} = 0.$$

The longitudinal component (parallel to the propagation direction) of the electric field is thus a constant. We choose it to be $E_x = 0$, since a constant different from zero does not represent a propagating wave. Similarly, $\nabla \cdot \mathbf{B} = 0$ implies $B_x = 0$. In conclusion, the electric and magnetic fields are transverse: they both belong to a plane Π perpendicular to the propagation axis $\mathbf{n} = \mathbf{u}_x$.



Hence,

$$\mathbf{E} \cdot \mathbf{n} = \mathbf{B} \cdot \mathbf{n} = 0 \quad (10.13)$$

and so,

$$\begin{aligned} \mathbf{E}(x, t) &= E_y(x, t)\mathbf{u}_y + E_z(x, t)\mathbf{u}_z, \\ \mathbf{B}(x, t) &= B_y(x, t)\mathbf{u}_y + B_z(x, t)\mathbf{u}_z. \end{aligned}$$

- STRUCTURE OF A PLANE WAVE

We now consider a progressive plane wave that propagates in the direction of increasing values of x . The general solution reads

$$\begin{aligned} \mathbf{E} &= E_y(x - ct)\mathbf{u}_y + E_z(x - ct)\mathbf{u}_z, \\ \mathbf{B} &= B_y(x - ct)\mathbf{u}_y + B_z(x - ct)\mathbf{u}_z. \end{aligned}$$

Considering Maxwell-Faraday's equation, $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$, projected along \mathbf{u}_x or \mathbf{u}_y , we obtain

$$\begin{aligned}\frac{\partial E_z(x - ct)}{\partial x} &= \frac{\partial B_y(x - ct)}{\partial t}, \\ \frac{\partial E_y(x - ct)}{\partial x} &= -\frac{\partial B_z(x - ct)}{\partial t}.\end{aligned}$$

and writing $X = x - ct$, these equations become

$$\begin{aligned}\frac{\partial E_z(X)}{\partial X} &= -c \frac{\partial B_y(X)}{\partial X}, \\ \frac{\partial E_y(X)}{\partial X} &= c \frac{\partial B_z(X)}{\partial X},\end{aligned}$$

and so, excluding constant solutions, we obtain

$$\begin{aligned}-E_z/c &= B_y, \\ E_y/c &= B_z,\end{aligned}$$

which implies that \mathbf{E} and \mathbf{B} are perpendicular to each other. Indeed,

$$\mathbf{E} \cdot \mathbf{B} = E_y B_y + E_z B_z = E_y(-E_z/c) + E_z(E_y/c) = 0.$$

Note also that

$$\mathbf{u}_x \times \mathbf{E} = E_y \mathbf{u}_z - E_z \mathbf{u}_y = c \mathbf{B}.$$

More generally, for a wave propagating in vacuum along the direction \mathbf{n} , we have

$$\mathbf{B} = \frac{1}{c} \mathbf{n} \times \mathbf{E}. \quad (10.14)$$

The vectors $(\mathbf{n}, \mathbf{E}, \mathbf{B})$ form therefore a direct trihedral.

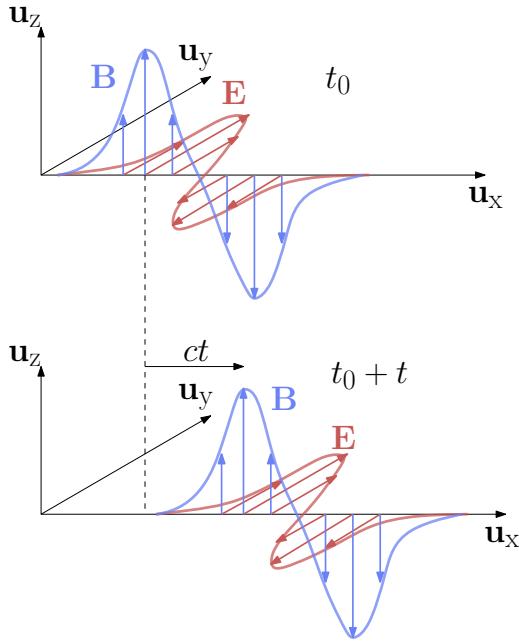
Note that for a regressive wave, the same result holds after changing \mathbf{n} into $-\mathbf{n}$.

10.3.5 Polarization

By convention, the polarization of the wave specifies the direction of the electric field. For a progressive wave propagating along the x -axis, we say that it is linearly polarized along the y -direction if the electric field is along \mathbf{u}_y at all times,

$$\begin{aligned}\mathbf{E} &= E_y(x - ct) \mathbf{u}_y, \\ \mathbf{B} &= B_z(x - ct) \mathbf{u}_z = \frac{1}{c} E_y(x - ct) \mathbf{u}_z,\end{aligned}$$

and such a linearly-polarized wave is illustrated in the figure below



Note also that any plane wave propagating along the x -axis can be written as a superposition of two independent plane waves according to Eq. (10.14),

$$\underbrace{\{E_y, B_z\}}_{\text{linear polarization along } y} \quad \underbrace{\{E_z, B_y\}}_{\text{linear polarization along } z}$$

Light polarization is discussed in more detail in section 11.11 of F. Cadiz's textbook.

• 1D SINUSOIDAL PLANE WAVE

As seen in section 10.2.2, a sinusoidal plane wave propagating along the x -axis can be written as

$$\mathbf{E} = \mathbf{E}_0 \cos(kx - \omega t + \phi_0),$$

where $\mathbf{E}_0 \perp \mathbf{u}_x$, and where the spatial and temporal oscillation frequencies of the wave are related by the dispersion relation

$$\omega = kc,$$

or, equivalently, the periods are related by

$$\lambda = cT,$$

where λ is called the wavelength. Visible light corresponds to electromagnetic waves in the range of wavelengths $400 \text{ nm} < \lambda < 800 \text{ nm}$. The wavelengths shorter than 400 nm correspond to ultraviolet radiation, while those longer than 800 nm are called infrared radiation. Table 10.1 provides a list of different waves of the electromagnetic spectrum according to their wavelength or frequency. The propagation of all these waves from radio-waves to γ -rays is governed by the same principles, summarized in Maxwell's equations.

Table 10.1: The electromagnetic spectrum

Name	λ	$f = \omega/2\pi$
Radio waves	10 cm to 10 km	10^4 to 10^9 Hz
Microwaves	1 mm to 10 cm	10^9 to 10^{12} Hz
Infrared	750 nm to 1 mm	10^{12} to 10^{14} Hz
Visible light	400 to 750 nm	$\sim 5 \times 10^{14}$ Hz
Ultraviolet (UV)	10 to 400 nm	10^{15} to 10^{16} Hz
X-rays	0.01 to 10 nm	10^{16} to 10^{19} Hz
γ rays	< 10 pm	$> 10^{19}$ Hz

10.4 Sinusoidal electromagnetic plane waves

Consider now a more general case in which a sinusoidal electromagnetic wave propagates along the unit direction \mathbf{n} , and let us choose an orthogonal basis $(\mathbf{u}, \mathbf{v}, \mathbf{n})$ such that $\mathbf{u} \times \mathbf{v} = \mathbf{n}$. The most general sinusoidal wave will have an electric field of the form

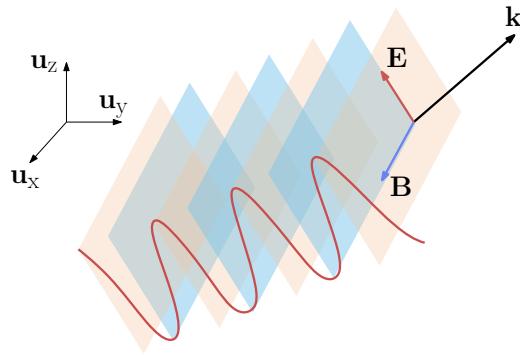
$$\mathbf{E} = E_1 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \phi_1) \mathbf{u} + E_2 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \phi_2) \mathbf{v},$$

where the wavevector is defined as

$$\mathbf{k} = k\mathbf{n} = k_x \mathbf{u}_x + k_y \mathbf{u}_y + k_z \mathbf{u}_z,$$

so that its direction coincides with the propagation direction of the wave. At a given instant t , the wavefronts are planes perpendicular to the wavevector \mathbf{k} such that the phase $\mathbf{k} \cdot \mathbf{x} + \phi$ is constant on them. These planes are therefore given by

$$\mathbf{k} \cdot \mathbf{x} = k_x x + k_y y + k_z z = \text{constant}.$$



In a complex representation, the electric field of such a wave can be rewritten as $\mathbf{E} = \text{Re}\{\underline{\mathbf{E}}\}$, with

$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

and

$$\underline{\mathbf{E}}_0 = E_1 e^{i\phi_1} \mathbf{u} + E_2 e^{i\phi_2} \mathbf{v}.$$

Injecting this wave in the 3D d'Alembert equation gives the dispersion relation

$$\omega = |\mathbf{k}|c = kc.$$

Finally, if we similarly write the magnetic field as $\mathbf{B} = \text{Re}\{\underline{\mathbf{B}}\} = \underline{\mathbf{B}}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$, the four Maxwell equations in vacuum for a sinusoidal plane wave become

$\mathbf{k} \cdot \underline{\mathbf{E}} = 0$	Gauss-Maxwell,
$\mathbf{k} \cdot \underline{\mathbf{B}} = 0$	Thomson-Maxwell,
$\mathbf{k} \times \underline{\mathbf{E}} = \omega \underline{\mathbf{B}}$	Faraday-Maxwell,
$\mathbf{k} \times \underline{\mathbf{B}} = -\frac{\omega}{c^2} \underline{\mathbf{E}}$	Ampère-Maxwell.

(10.15)

Note that the Gauss-Maxwell and Thomson-Maxwell equations for a sinusoidal progressive plane wave imply the transversality of both the electric and magnetic fields

$$\begin{aligned} \mathbf{k} \cdot \underline{\mathbf{E}} = 0 &\rightarrow \underline{\mathbf{E}} \perp \mathbf{k}, \\ \mathbf{k} \cdot \underline{\mathbf{B}} = 0 &\rightarrow \underline{\mathbf{B}} \perp \mathbf{k}. \end{aligned}$$

Since $\mathbf{k} = k\mathbf{n}$, Faraday-Maxwell's law becomes

$$\mathbf{k} \times \underline{\mathbf{E}} = \omega \underline{\mathbf{B}} \rightarrow \underline{\mathbf{B}} = \frac{k\mathbf{n}}{\omega} \times \underline{\mathbf{E}} \quad (10.16)$$

and using the dispersion relation $k/\omega = 1/c$, we retrieve the previously found relationship between the magnetic and the electric field, now in the particular case of a sinusoidal plane wave,

$$\underline{\mathbf{B}} = \frac{1}{c} \mathbf{n} \times \underline{\mathbf{E}}. \quad (10.17)$$

10.5 Electromagnetic energy and Poynting's theorem

Not only Maxwell's equations contain the conservation of charge in them, but it is also possible to obtain from Maxwell's equations an equation reflecting the conservation of electromagnetic energy, called Poynting's theorem. In its integral form, such an equation should read

$$\frac{dU(t)}{dt} = \iiint_{\Omega} \frac{\partial u_{EM}(\mathbf{x}, t)}{\partial t} d^3x = -W_1 - W_2.$$

where

$$U(t) = \iiint_{\Omega} u_{EM}(\mathbf{x}, t) d^3x$$

is the total electromagnetic energy contained in an arbitrary volume $\Omega \subseteq \mathbb{R}^3$, and $u_{EM}(\mathbf{x}, t)$ is the density of electromagnetic energy per unit volume (J m^{-3}). This equation states that the variation of

the total energy contained in Ω between t and $t + dt$ may have two origins: an energy exchange with the charges in Ω (at a rate W_1) or a transfer of energy through the frontier $\partial\Omega$ (at a rate W_2). The term W_1 is simply the work done by the Lorentz force on the charges contained in Ω : a fraction of the electromagnetic energy of the wave is thus converted into kinetic energy of the charges. We calculate first the Lorentz force exerted on a volume element d^3x due to the passage of the wave. If the velocity of the charges in d^3x is \mathbf{v} , then,

$$d\mathbf{f} = dq(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \rho d^3x(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

Then, the work done by the Lorentz force on Ω reads

$$W_1 = \iiint_{\Omega} d\mathbf{f} \cdot \mathbf{v} = \iiint_{\Omega} \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} d^3x = \iiint_{\Omega} \rho \mathbf{v} \cdot \mathbf{E} d^3x$$

and since $\rho \mathbf{v} = \mathbf{J}$, where \mathbf{J} is the current density,

$$W_1 = \iiint_{\Omega} \mathbf{J} \cdot \mathbf{E} d^3x.$$

For the term W_2 , we assume that the transfer of energy through the surface $\partial\Omega$ can be written in the form of a flux through the latter,

$$W_2 = \Phi_{\partial\Omega, \mathbf{\Pi}} = \oint_{\partial\Omega} \mathbf{\Pi} \cdot d\mathbf{S} = \iiint_{\Omega} \nabla \cdot \mathbf{\Pi} d^3x,$$

where $\mathbf{\Pi}$, the Poynting vector, represents a flux of energy associated with the electromagnetic wave, i.e., a density of energy per unit surface and per unit time (W/m^2). The flux of $\mathbf{\Pi}$ through $\partial\Omega$ represents the power transmitted by the electromagnetic wave through the surface. In this way, the conservation equation for the total electromagnetic energy reads

$$\iiint_{\Omega} \frac{\partial u_{EM}(\mathbf{x}, t)}{\partial t} d^3x = - \iiint_{\Omega} \mathbf{J} \cdot \mathbf{E} d^3x - \iiint_{\Omega} \nabla \cdot \mathbf{\Pi} d^3x,$$

or, equivalently, in its local form, Poynting's theorem reads

$$\frac{\partial u_{EM}}{\partial t} + \nabla \cdot \mathbf{\Pi} = -\mathbf{J} \cdot \mathbf{E}. \quad (10.18)$$

10.5.1 Poynting vector and density of electromagnetic energy

We will now use Maxwell's equations in order to demonstrate Eq. (10.18) and identify $\mathbf{\Pi}$ and u_{EM} in terms of the fields \mathbf{E} and \mathbf{B} . For this we start with Ampère-Maxwell's law in the form, $\nabla \times \frac{\mathbf{B}}{\mu_0} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$, and take the scalar product with \mathbf{E} on both sides of the equation to find

$$\frac{1}{\mu_0} \mathbf{E} \cdot (\nabla \times \mathbf{B}) = \mathbf{J} \cdot \mathbf{E} + \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J} \cdot \mathbf{E} + \frac{\partial}{\partial t} \left(\frac{\epsilon_0 \mathbf{E}^2}{2} \right). \quad (10.19)$$

Similarly, we use Maxell-Faraday's law, $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$, and we take the scalar product with \mathbf{B}/μ_0 to find

$$\frac{1}{\mu_0} \mathbf{B} \cdot (\nabla \times \mathbf{E}) = -\frac{\mathbf{B}}{\mu_0} \cdot \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t} \left(\frac{\mathbf{B}^2}{2\mu_0} \right). \quad (10.20)$$

Subtracting Eq. (10.20) from Eq. (10.19), we obtain

$$\frac{1}{\mu_0} (\mathbf{E} \cdot (\nabla \times \mathbf{B}) - \mathbf{B} \cdot (\nabla \times \mathbf{E})) = \mathbf{J} \cdot \mathbf{E} + \frac{\partial}{\partial t} \left(\frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} \right). \quad (10.21)$$

Finally, using the identity

$$\mathbf{E} \cdot (\nabla \times \mathbf{B}) - \mathbf{B} \cdot (\nabla \times \mathbf{E}) = -\nabla \cdot (\mathbf{E} \times \mathbf{B}),$$

Eq. (10.21) is rewritten as

$$-\frac{1}{\mu_0} \nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{J} \cdot \mathbf{E} + \frac{\partial}{\partial t} \left(\frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} \right)$$

and we recognize an equation of the form of Eq. (10.18).

Definition 10.3: Poynting's theorem

The conservation of the electromagnetic energy in its local form is given by Poynting's theorem,

$$\nabla \cdot \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \right) + \frac{\partial}{\partial t} \left(\frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} \right) = -\mathbf{J} \cdot \mathbf{E}, \quad (10.22)$$

or, equivalently,

$$\frac{\partial}{\partial t} u_{EM} + \nabla \cdot \mathbf{\Pi} = -\mathbf{J} \cdot \mathbf{E}, \quad (10.23)$$

where $\mathbf{\Pi}$ is the Poynting vector, which gives the direction of electromagnetic energy flow,

$$\mathbf{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \quad (10.24)$$

and u_{EM} the density of the electromagnetic energy,

$$u_{EM} = \frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0}. \quad (10.25)$$

10.5.2 Poynting vector of a sinusoidal plane wave

Let us apply Poynting's theorem to a sinusoidal, plane electromagnetic wave in vacuum propagating along the x -axis and linearly-polarized along \mathbf{u}_y . The electric and magnetic fields read

$$\begin{aligned} \mathbf{E} &= E_0 \cos(\omega t - kx) \mathbf{u}_y, \\ \mathbf{B} &= \frac{1}{c} \mathbf{u}_x \times \mathbf{E} = \frac{1}{c} E_0 \cos(\omega t - kx) \mathbf{u}_z. \end{aligned}$$

The electromagnetic energy density associated with such a wave is equally distributed between its electric and magnetic components, as shown by the calculation

$$u_{EM} = \frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} = \frac{\epsilon_0}{2} E_0^2 \cos^2(\omega t - kx) + \frac{E_0^2}{2c^2\mu_0} \cos^2(\omega t - kx) = \epsilon_0 E_0^2 \cos^2(\omega t - kx),$$

where the relation $\epsilon_0\mu_0 c^2 = 1$ was used, and the Poynting vector reads

$$\boldsymbol{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} = c\epsilon_0 E_0^2 \cos^2(\omega t - kx) \mathbf{u}_x = cu_{EM} \mathbf{u}_x.$$

This result means that the electromagnetic energy is moving at velocity c in the positive x direction. Indeed, the Poynting vector represents the energy passing per unit area and per unit time through a surface element perpendicular to $\boldsymbol{\Pi}$. Note that both u_{EM} and $\boldsymbol{\Pi}$ oscillate at the frequency ω , which in the visible range (600 nm) corresponds to a period $T = 2\pi/\omega$ of the order of $T \sim 2$ fs, which is too fast to be detected even with our fastest optical detectors. As an example, the eye response time is of the order of 0.1 s, a CCD camera responds typically within 10 ms and a fast photodiode or a photomultiplier may achieve a response of 1 ns. Any light detector thus measure an averaged power over the detector response time $T_{\text{det}} \gg T$. The time average of the electromagnetic energy density and Poynting's vector for a sinusoidal plane wave read

$$\begin{aligned} \langle u_{EM} \rangle_{T_{\text{det}}} &= \epsilon_0 E_0^2 \frac{1}{T_{\text{det}}} \int_0^{T_{\text{det}}} \underbrace{\cos^2(\omega t - kx) dt}_{=\frac{1}{2}(1+\cos(2(\omega t - kx)))} = \epsilon_0 E_0^2 \left(\frac{1}{2} + \frac{T}{4\pi T_{\text{det}}} [\sin 2(\omega t - kx)]_0^{T_{\text{det}}} \right) \approx \frac{\epsilon_0 E_0^2}{2}, \end{aligned}$$

since $T_{\text{det}} \gg T$. Note therefore that for a sinusoidal wave, the detector measures an average energy equal to the average over one period,

$$\langle u_{EM} \rangle_T = \epsilon_0 E_0^2 \underbrace{\frac{1}{T} \int_0^T \cos^2(\omega t - kx) dt}_{=1/2} = \frac{\epsilon_0 E_0^2}{2}$$

and

$$\langle \boldsymbol{\Pi} \rangle_T = c\epsilon_0 E_0^2 \frac{1}{T} \int_0^T \cos^2(\omega t - kx) dt \mathbf{u}_x = \frac{1}{2} c\epsilon_0 E_0^2 \mathbf{u}_x = c \langle u_{EM} \rangle_T \mathbf{u}_x.$$

Remarks

1. Since the Poynting vector is not a linear function of the fields, we cannot calculate it by using the complex representation directly, since

$$\boldsymbol{\Pi} \neq \text{Re} \left\{ \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \right\}$$

however, for the time-averaged energy density and time-averaged Poynting vector, we can write:

$$\langle u_{EM} \rangle_T = \frac{1}{2} \text{Re} \left\{ \frac{\epsilon_0}{2} \underline{\mathbf{E}} \cdot \underline{\mathbf{E}}^* + \frac{1}{2\mu_0} \underline{\mathbf{B}} \cdot \underline{\mathbf{B}}^* \right\}$$

and define the complex Poynting vector

$$\underline{\boldsymbol{\Pi}} = \frac{\underline{\mathbf{E}} \times \underline{\mathbf{B}}^*}{2\mu_0},$$

such that

$$\langle \underline{\Pi} \rangle_T = \operatorname{Re}\{\underline{\Pi}\} = \frac{1}{2\mu_0} \operatorname{Re}\{\underline{\mathbf{E}} \times \underline{\mathbf{B}}^*\}. \quad (10.26)$$

2. The electromagnetic energy density of a plane wave is uniform, so that the plane wave, which is infinite in all directions and in time, carries an infinite electromagnetic energy. Whereas this is not physically acceptable, one must recall that plane waves are useful because they can approximate real waves locally, and because every real wave is a linear superposition of plane waves.

Definition 10.4: Intensity of an electromagnetic wave

The intensity \mathcal{I} , also called irradiance, of an electromagnetic wave corresponds to the average power going through a unit surface perpendicular to the propagation axis. It has units of W m^{-2} and corresponds to the time-average of the modulus of the Poynting vector over the detector response time T_{det} which, for sinusoidal plane waves and for $T_{\text{det}} \gg T$, coincides with the average over one period T of the wave oscillation,

$$\mathcal{I} = \langle |\underline{\Pi}| \rangle_T = \left\langle \left| \frac{\underline{\mathbf{E}} \times \underline{\mathbf{B}}}{\mu_0} \right| \right\rangle_T. \quad (10.27)$$

For a sinusoidal plane wave, it reads

$$\mathcal{I} = \frac{c\epsilon_0}{2} E_0^2, \quad (10.28)$$

where E_0 is the amplitude of the wave electric field. Every optical detector of detection surface S , oriented perpendicular to the propagation axis, measures an average power given by $\mathcal{P} = \mathcal{I}S \propto E_0^2$.

Suggested Exercises: Ex 11.5, 11.6, 11.7 in F. Cadiz's textbook.

10.6 Electromagnetic waves in matter

In this section we will consider the propagation of electromagnetic waves in matter (dielectric media, magnetized media, metallic media). Recall that a dielectric medium is characterized by the polarization density vector \mathbf{P} which corresponds to the average dipole moment per unit volume. The polarization charge density reads $\rho_P = -\nabla \cdot \mathbf{P}$ and must be added to the free charge density ρ in Gauss's law for the spatially averaged electric field

$$\nabla \cdot \mathbf{E} = \frac{\rho + \rho_P}{\epsilon_0} = \frac{\rho}{\epsilon_0} - \frac{1}{\epsilon_0} \nabla \cdot \mathbf{P}.$$

Defining the electric displacement field, $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, Gauss's law in a dielectric reads

$$\nabla \cdot \mathbf{D} = \rho.$$

Similarly, a magnetized media is characterized by the magnetization vector \mathbf{M} which corresponds to the average magnetic moment per unit volume. The magnetization makes a contribution to the current density via the magnetization current density $\mathbf{J}_M = \nabla \times \mathbf{M}$ that must be added to the free current density, \mathbf{J} so that Ampère's law for the average magnetic field reads

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \mathbf{J}_M = \mu_0 \mathbf{J} + \mu_0 \nabla \times \mathbf{M}.$$

Defining the auxiliary magnetic field $\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}$, Ampère's law can be rewritten as

$$\nabla \times \mathbf{H} = \mathbf{J}.$$

Now we know that in vacuum, Ampère's law is incomplete and that we must include the displacement current, $\mathbf{J}_D = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$, to the right-hand side in order to ensure charge conservation in the time-dependent regime. Thus, we could expect that Ampère-Maxwell's equation in matter and in the time dependent regime will read

$$\nabla \times \mathbf{H} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

However, the latter equation is not yet complete. The expression for displacement current in the case of vacuum was inferred from the requirement that the total current be divergence-free ($\nabla \cdot (\mathbf{J} + \mathbf{J}_D) = 0$). Since we take into account polarization charges in a dielectric medium, charge conservation must involve polarization charges as well. In the time-dependent regime, this will give rise to polarization current \mathbf{J}_P that must be taken into account in Ampère-Maxwell's equation.

10.6.1 General conservation principle

The conservation principle for any quantity states that the variation of a quantity $Q = \iiint_{\Omega} \rho_Q(\mathbf{x}, \mathbf{t}) d^3x$ within a certain volume Ω , where $\rho_Q(\mathbf{x}, t)$ denotes the density of Q , is contributed by the incoming flux of an associated current transporting the quantity through the boundary $\partial\Omega$, and the net creation of this quantity (difference between generation and losses) within the volume. In other words, the conservation equation for any quantity Q is expressed by the change in time

$$\frac{dQ}{dt} = \iiint_{\Omega} \frac{\partial \rho_Q(\mathbf{x}, \mathbf{t})}{\partial t} d^3x = - \iint_{\partial\Omega} \mathbf{J}_Q \cdot \mathbf{n} dS + \iiint_{\Omega} w_{\text{gen}}(\mathbf{x}, \mathbf{t}) d^3x - \iiint_{\Omega} w_{\text{loss}}(\mathbf{x}, \mathbf{t}) d^3x,$$

where $w_{\text{gen}}(\mathbf{x}, \mathbf{t})$ and $w_{\text{loss}}(\mathbf{x}, \mathbf{t})$ denote generation and loss rates for the quantity Q , which are non-zero if there are sources or sinks within Ω .

Ostrogradsky's theorem allows us to transform the flux into a volume integral

$$\iint_{\partial\Omega} \mathbf{J}_Q \cdot \mathbf{n} dS = \iiint_{\Omega} \nabla \cdot \mathbf{J}_Q d^3x.$$

Since this is valid for any volume Ω , it applies locally, that is, for an infinitesimal volume $\Omega \equiv d^3x$ centered at \mathbf{x} , hence, the local conservation equation reads

$$\frac{\partial \rho_Q}{\partial t} + \nabla \cdot \mathbf{J}_Q = w_{\text{gen}} - w_{\text{loss}}.$$

The definition of the current \mathbf{J}_Q and rates w_{gen} , w_{loss} , are specific to the quantity Q . Below, we apply this principle to the charge and to the electromagnetic energy.

10.6.2 Displacement current and Maxwell's equations in matter

In a time-dependent regime, the polarization density $\mathbf{P}(\mathbf{x}, t)$ in a dielectric will depend on time in response to a time-varying electric field $\mathbf{E}(\mathbf{x}, t)$. The oscillation of the polarization charges will also contribute to the total current density via the polarization current density \mathbf{J}_P , whose expression is obtained from the conservation equation of the polarization charge.

The polarization charge density is defined in the same way as in the permanent regime, that is,

$$\rho_P(\mathbf{x}, t) = -\nabla \cdot \mathbf{P}(\mathbf{x}, t).$$

Deriving in time, we find

$$\frac{\partial \rho_P}{\partial t} = -\frac{\partial (\nabla \cdot \mathbf{P})}{\partial t} = -\nabla \cdot \frac{\partial \mathbf{P}}{\partial t}.$$

We see that this equation takes the form of a general conservation equation (no generation or losses)

$$\frac{\partial \rho_P}{\partial t} + \nabla \cdot \mathbf{J}_P = 0$$

where the polarization current density reads

$$\mathbf{J}_P = \frac{\partial \mathbf{P}}{\partial t}. \quad (10.29)$$

From here, two equivalent paths can be followed to generalize Ampère-Maxwell's equation for matter to the time-dependent regime.

- Starting from Ampère-Maxwell's equation in matter and in the permanent regime, the polarization current can be simply added to the free current density \mathbf{J} *and* to the displacement current $\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$ that was already inferred in the case of vacuum, so that in a magnetized dielectric medium, the complete equation becomes

$$\nabla \times \mathbf{H} = \mathbf{J} + \underbrace{\frac{\partial \mathbf{P}}{\partial t}}_{\mathbf{J}_P} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

and recalling that $\mathbf{P} + \epsilon_0 \mathbf{E} = \mathbf{D}$, Ampère-Maxwell's equation in matter finally reads

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \quad (10.30)$$

Note that Eq. (10.30) is compatible with the conservation law of conduction (or free) charge, since

$$\nabla \cdot (\nabla \times \mathbf{H}) = 0 = \nabla \cdot \mathbf{J} + \frac{\partial}{\partial t} \nabla \cdot \mathbf{D}$$

and from Gauss's law, $\nabla \cdot \mathbf{D} = \rho$, we retrieve

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

- Alternatively, we can reproduce the same approach as that followed for vacuum. Starting from Ampère-Maxwell's equation in matter for permanent regime, $\nabla \times \mathbf{H} = \mathbf{J}$, we can add a displacement current \mathbf{J}_D to the conduction current \mathbf{J} and find its expression by requiring that charge conservation be satisfied. Hence, postulating that in the time-dependent regime Ampère-Maxwell's equation reads

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{J}_D,$$

and taking the divergence on both sides, the divergence of the left-hand side is identically zero ($\nabla \cdot (\nabla \times \mathbf{H}) = 0$), thus

$$\nabla \cdot (\mathbf{J} + \mathbf{J}_D) = 0.$$

The conduction current satisfies charge conservation $\partial_t \rho + \nabla \cdot \mathbf{J} = 0$, and Gauss-Maxwell's equation allows us to express the density of conduction charges as a function of the displacement field, $\rho = \nabla \cdot \mathbf{D}$, leading to

$$-\nabla \cdot \mathbf{J}_D = \nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} = -\frac{\partial(\nabla \cdot \mathbf{D})}{\partial t} = -\nabla \cdot \frac{\partial \mathbf{D}}{\partial t}.$$

This means that the displacement current that is compatible with the conservation law for conduction charges is the time derivative of the displacement field

$$\boxed{\mathbf{J}_D = \frac{\partial \mathbf{D}}{\partial t}},$$

and Ampère-Maxwell's equation in matter then reads

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t},$$

that is, this path leads to Eq. (10.30) as expected.

In a dielectric, the displacement current $\mathbf{J}_D = \frac{\partial \mathbf{D}}{\partial t}$ can be written as the superposition of the real current polarization current \mathbf{J}_P , associated with the movement of bound charges and the displacement current in vacuum ($\epsilon_0 \partial \mathbf{E} / \partial t$). This results from the definition of the displacement field and polarization current

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad \text{and} \quad \mathbf{J}_P = \frac{\partial \mathbf{P}}{\partial t},$$

leading to

$$\mathbf{J}_D = \frac{\partial \mathbf{D}}{\partial t} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_P.$$

An expansion of the auxiliary fields in Ampère-Maxwell's equation then reads

$$\nabla \times \underbrace{\mathbf{H}}_{\mathbf{B}/\mu_0 - \mathbf{M}} = \mathbf{J} + \frac{\partial}{\partial t} \underbrace{\mathbf{D}}_{\epsilon_0 \mathbf{E} + \mathbf{P}},$$

leading to

$$\nabla \times \frac{\mathbf{B}}{\mu_0} = \underbrace{\nabla \times \mathbf{M}}_{\mathbf{J}_M} + \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \underbrace{\frac{\partial \mathbf{P}}{\partial t}}_{\mathbf{J}_P},$$

that is,

$$\nabla \times \mathbf{B} = \mu_0 (\mathbf{J} + \mathbf{J}_M + \mathbf{J}_P) + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}.$$

The latter form of Ampère-Maxwell's equation is similar to its counterpart in vacuum but the current includes three contributions, the conduction current \mathbf{J} , the magnetization current \mathbf{J}_M and the polarization current \mathbf{J}_P . In other words, in the time-dependent as well as in the permanent regime, Maxwell's equation in matter can be written either using the auxiliary fields \mathbf{D} and \mathbf{H} , or keeping the electric and magnetic fields \mathbf{E} and \mathbf{B} but adding polarization and magnetization charges and currents in source terms.

Definition 10.5: Maxwell's equations in matter

Maxwell's equations in matter are given by

$\nabla \cdot \mathbf{D} = \rho$	Gauss-Maxwell,
$\nabla \cdot \mathbf{B} = 0$	Thomson-Maxwell,
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday-Maxwell,
$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$	Ampère-Maxwell.

(10.31)

where ρ and \mathbf{J} denote the conduction (free) charge and current densities and the auxiliary fields are given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P},$$

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}.$$

The behavior of the different types of materials is determined by the constitutive relations

$$\mathbf{P}(\mathbf{E}) \quad \text{and} \quad \mathbf{M}(\mathbf{H}).$$

- From now on we will restrict ourselves to the propagation in linear, homogeneous and isotropic materials,

$$\mathbf{D} = \epsilon \mathbf{E} \quad \text{and} \quad \mathbf{B} = \mu \mathbf{H}$$

where the permittivity ϵ and the permeability μ are scalar quantities. In this case, The Gauss-Maxwell and Ampère-Maxwell equations take the same form as in vacuum provided the permittivity and the permeability are replaced by their counterparts in matter:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon},$$

$$\nabla \times \mathbf{B} = \mu \mathbf{J} + \epsilon \mu \frac{\partial \mathbf{E}}{\partial t}.$$

- For non magnetic materials, $\mathbf{M} = 0$ so that $\mathbf{B} = \mu_0 \mathbf{H}$ and $\mu = \mu_0$.
- A conductor is characterized by its conductivity. According to Ohm's law,

$$\mathbf{J} = \sigma \mathbf{E}.$$

- Remark: these examples of constitutive relations were established in the permanent regime (electrostatics, magnetostatics). Assuming their validity in a time-dependent regime implies an instantaneous response of the medium to the variations of the electric field. In reality, the inertia and friction of charges in matter imposes a delay between the cause (**E**) and the effects (**D** and **J**). This will be discussed in the next chapter.

10.6.3 Poynting's theorem in matter

From Maxwell's equations in matter (10.31), it is possible to generalize Poynting's theorem that was previously demonstrated for the electromagnetic fields in vacuum. Taking the scalar product of Ampère-Maxwell's law with **E** on both sides of the equation, we obtain

$$\mathbf{E} \cdot (\nabla \times \mathbf{H}) = \mathbf{J} \cdot \mathbf{E} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}. \quad (10.32)$$

Similarly, we use Maxell-Faraday's law and we take the scalar product with **H** to obtain

$$\mathbf{H} \cdot (\nabla \times \mathbf{E}) = -\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}. \quad (10.33)$$

Subtracting Eq. (10.33) from Eq. (10.32), we obtain

$$\underbrace{\mathbf{E} \cdot (\nabla \times \mathbf{H}) - \mathbf{H} \cdot (\nabla \times \mathbf{E})}_{-\nabla \cdot (\mathbf{E} \times \mathbf{H})} = \mathbf{J} \cdot \mathbf{E} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \quad (10.34)$$

and we recognize an equation which is of the form of a local conservation equation

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{J} \cdot \mathbf{E},$$

if we remember the expressions for the differential of the density of electric and magnetic free energy

$$df = \underbrace{\mathbf{E} \cdot d\mathbf{D}}_{\text{electric}} + \underbrace{\mathbf{H} \cdot d\mathbf{B}}_{\text{magnetic}}.$$

Now for a linear, homogeneous and isotropic medium, the equations of state $\mathbf{D} = \epsilon \mathbf{E}$ and $\mathbf{B} = \mu \mathbf{H}$ allow us to recognize that the free energy density is nothing but the density of electromagnetic energy

$$f = u_{EM} = \epsilon \frac{\mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu}.$$

Integrating

$$df = \epsilon \mathbf{E} \cdot d\mathbf{E} + \mu \mathbf{H} \cdot d\mathbf{H},$$

we find indeed

$$f = \epsilon \frac{\mathbf{E}^2}{2} + \mu \frac{\mathbf{H}^2}{2} = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}),$$

and we retrieve a local conservation equation similar to that of Eq. (10.18) for the electromagnetic energy in vacuum.

Definition 10.6: Poynting's theorem in matter

Poynting's theorem in matter reads

$$\frac{\partial}{\partial t} u_{EM} + \nabla \cdot \mathbf{\Pi} = -\mathbf{J} \cdot \mathbf{E}, \quad (10.35)$$

where the Poynting's vector is given by

$$\mathbf{\Pi} = \mathbf{E} \times \mathbf{H} \quad (10.36)$$

and the density of electromagnetic energy for a linear, homogeneous and isotropic medium is

$$u_{EM} = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}).$$

10.6.4 Propagation of waves in dielectric media

Let us now consider the case of a linear homogeneous and isotropic dielectric medium in the absence of free charges. Assuming that we may write $\mathbf{D}(\mathbf{x}, t) = \epsilon \mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t) = \mu \mathbf{H}(\mathbf{x}, t)$, it is easy to verify that the propagation equation for the electric and magnetic field becomes

$$\begin{aligned} \nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} &= 0, \\ \nabla^2 \mathbf{B} - \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} &= 0. \end{aligned} \quad (10.37)$$

There are two possible ways to demonstrate Eq. (10.37):

- **Directly from Maxwell's equations:** We will keep the conduction charge and current for this demonstration. As in vacuum, we first take the curl of Faraday's law and commute the curl with differentiation with respect to time. We find

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t} \nabla \times \mathbf{B}.$$

Then, on the left-hand side, we expand the double curl using the identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and on the right-hand side, we introduce Ampère-Maxwell's equation $\nabla \times \mathbf{B} = \mu(\mathbf{J} + \epsilon \partial \mathbf{E} / \partial t)$. We find

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{\partial}{\partial t} \mu \left(\mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right).$$

In the first and last terms, we use Gauss-Maxwell's equation, leading to $\nabla(\nabla \cdot \mathbf{E}) = \nabla \rho / \epsilon$. Hence,

$$\nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\nabla \rho}{\epsilon} + \mu \frac{\partial \mathbf{J}}{\partial t}. \quad (10.38)$$

This is the propagation equation for the electric field in matter, with source terms on the right-hand side, which coincides with the first of Eqs (10.37) in the absence of conduction charge and current.

For the magnetic field, take the curl of Ampère-Maxwell's equation:

$$\nabla \times (\nabla \times \mathbf{B}) = \mu \nabla \times \mathbf{J} + \mu\epsilon \nabla \times \frac{\partial \mathbf{E}}{\partial t}.$$

Then, on the left-hand side, expand the double curl as $\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}$, eliminate $\nabla \cdot \mathbf{B}$ (zero from Thomson-Maxwell's equation), commute the differentiation in time and the curl in the last term, and use Faraday's law $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$. This leads to

$$\nabla^2 \mathbf{B} - \epsilon\mu \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu \nabla \times \mathbf{J}, \quad (10.39)$$

which coincides with the second of Eqs (10.37) in the absence of conduction current.

- **From scalar and vector potentials:** In matter, a consequence of Thomson-Maxwell's equation is that the magnetic field writes as $\mathbf{B} = \nabla \times \mathbf{A}$. The electric field is also linked to the scalar and vector potentials by the same relation as in vacuum, $\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}$. Ampère-Maxwell's equation can then be rewritten using the potentials instead of the fields:

$$\nabla \times (\nabla \times \mathbf{A}) = \mu \mathbf{J} - \epsilon\mu \frac{\partial}{\partial t} \left(\nabla V + \frac{\partial \mathbf{A}}{\partial t} \right).$$

Then use the double curl identity $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ and collect the terms that form a d'Alembert equation on the left-hand side. On the right hand side, two terms appear in the form of a gradient, after commuting spatial and temporal differentiation:

$$\nabla^2 \mathbf{A} - \epsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{J} + \nabla \left(\epsilon\mu \frac{\partial V}{\partial t} + \nabla \cdot \mathbf{A} \right).$$

Now remember that the potentials V and \mathbf{A} are not uniquely defined. It is always possible to add the gradient of a scalar quantity $\phi(\mathbf{x}, t)$ to the vector potential and subtract the time-derivative of this quantity from the scalar potential to define exactly the same electric and magnetic fields. This means that an additional condition can be required to define the potentials. This is called a *gauge condition*. A convenient choice is the *Lorentz gauge*, which reads

$$\epsilon\mu \frac{\partial V}{\partial t} + \nabla \cdot \mathbf{A} = 0.$$

With this condition, the vector potential satisfies the propagation equation

$$\nabla^2 \mathbf{A} - \epsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{J}. \quad (10.40)$$

The scalar potential also satisfies a propagation equation that can be obtained from Gauss-Maxwell's equation

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon} \rightarrow \nabla \cdot \left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} \right) = \frac{\rho}{\epsilon} \rightarrow \nabla^2 V + \nabla \cdot \frac{\partial \mathbf{A}}{\partial t} = -\frac{\rho}{\epsilon}.$$

By differentiating the Lorentz gauge with respect to time, we find

$$\epsilon\mu \frac{\partial^2 V}{\partial t^2} = -\frac{\partial \nabla \cdot \mathbf{A}}{\partial t} = -\nabla \cdot \frac{\partial \mathbf{A}}{\partial t},$$

which allows us to write the propagation equation for the scalar potential as a d'Alembert equation

$$\nabla^2 V - \epsilon\mu \frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\epsilon}. \quad (10.41)$$

Equations (10.41) and (10.40) both take the form of d'Alembert's equation with source terms. Taking the curl of (10.40) leads to the propagation equation for the magnetic field Eq. (10.39). The propagation equation (10.38) for the electric field $\mathbf{E} = -\nabla V - \partial_t \mathbf{A}$, can in turn be retrieved by differentiating in time Eq. (10.40) and adding the gradient of Eq. (10.41), which completes the second way of deriving propagation equations for \mathbf{E} and \mathbf{B} in matter.

As a general comment on the propagation equations for the potentials, we note that in the permanent regime, we retrieve Poisson's equations for the vector and scalar potentials. In the absence of conduction charge ($\rho = 0$) or current ($\mathbf{J} = \mathbf{0}$), and in the time-dependent regime, propagation equations in matter for the electric and magnetic fields as well as for the scalar and vector potentials take the same form of d'Alembert's equation (10.37).

10.6.5 Resolution of propagation equations in matter

We limit the discussion here to the case of propagation in a (linear) material without conduction charge ($\rho = 0$) or current ($\mathbf{J} = \mathbf{0}$). Since propagation equations for \mathbf{E} , \mathbf{B} , \mathbf{A} and V are d'Alembert's equation, we can conclude that electromagnetic waves propagate in a dielectric just like in vacuum, except that the propagation velocity is smaller, given by

$$v = \frac{1}{\sqrt{\epsilon\mu}} = \frac{c}{\sqrt{\epsilon_r\mu_r}}.$$

All the properties discussed so far for electromagnetic waves in vacuum still hold if we replace the speed of light by its counterpart in the medium. In particular, the electric field and magnetic fields are transverse, that is, perpendicular to the propagation direction and \mathbf{E} , \mathbf{B} and the propagation direction \mathbf{n} form a direct trihedral with $\mathbf{B} = \frac{1}{v}\mathbf{n} \times \mathbf{E}$.

While results are similar to the case of vacuum, it is instructive to see an alternative derivation of these properties, based on a direct use of Maxwell's equations rather than d'Alembert's equation, which we develop below (and can be skipped in a first reading).

We will look for solutions in the form of plane waves, not necessarily sinusoidal, propagating along the x -axis of a Cartesian basis (\mathbf{u}_x , \mathbf{u}_y , \mathbf{u}_z). This means that the vector potential $\mathbf{A}(x, t)$ is independent of y , z , and so are the electric and magnetic fields $\mathbf{E}(x, t)$ and $\mathbf{B}(x, t)$. We can thus write the vector potential as

$$\mathbf{A}(x, t) = A_x(x, t)\mathbf{u}_x + A_y(x, t)\mathbf{u}_y + A_z(x, t)\mathbf{u}_z.$$

For a LHI medium without conduction charge or current, it is always possible to chose the scalar potential $V(x, t) = 0$ and this choice implies that the Lorentz gauge condition requires the vector potential to satisfy $\nabla \cdot \mathbf{A} = 0$. Hence, we have the following equations must be satisfied:

$$\begin{aligned}\text{Gauge condition: } \nabla \cdot \mathbf{A} &= 0 & \partial A_x / \partial x &= 0, \\ \text{Gauss-Maxwell: } \nabla \cdot \mathbf{E} &= 0 & \rightarrow \partial E_x / \partial x &= 0, \\ \text{Thomson-Maxwell: } \nabla \cdot \mathbf{B} &= 0 & \partial B_x / \partial x &= 0,\end{aligned}$$

which requires that A_x , E_x and B_x be independent of x , and thus only depend on time.

Next we use the two Maxwell equations involving a *curl*, which were the essential equations used to establish propagation equations. While developing the curl operator in a Cartesian basis, we directly set to zero the y - and z -derivatives as fields only depend on x and t .

$$\begin{aligned}\nabla \times \mathbf{E} = -\partial_t \mathbf{B} &\rightarrow \begin{vmatrix} \partial_x & E_x & -\partial_t B_x & \partial_t B_x = 0 \\ 0 & E_y & -\partial_t B_y & \partial_t B_y = \partial_x E_z \\ 0 & E_z & -\partial_t B_z & \partial_t B_z = -\partial_x E_y \end{vmatrix} \\ \nabla \times \mathbf{B} = v^{-2} \partial_t \mathbf{E} &\rightarrow \begin{vmatrix} \partial_x & B_x & \frac{1}{v^2} \partial_t E_x & \partial_t E_x = 0 \\ 0 & B_y & \partial_t E_y & -\partial_x B_z = v^{-2} \partial_t E_y \\ 0 & B_z & \partial_t E_z & \partial_x B_y = v^{-2} \partial_t E_z \end{vmatrix}\end{aligned}$$

The first equations in these two sets tell us that B_x and E_x do not depend on time. Since we already know that they do not depend on x , we conclude that $B_x = 0$ and $E_x = 0$ because we are not considering the possibility of static (constant) fields². This leads us to conclude that the electric and magnetic fields are transverse

$$\mathbf{E} \perp \mathbf{u}_x, \quad \mathbf{B} \perp \mathbf{u}_x.$$

The definition of the electric field from potentials $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$ together with $V = 0$ implies that $-\partial_t A_x = E_x = 0$, hence, A_x does not depend on time. The x -component of the vector potential is thus $A_x = 0$, meaning that \mathbf{A} is also transverse:

$$\mathbf{A} \perp \mathbf{u}_x$$

Writing $\mathbf{A}(x, t) = A_y(x, t)\mathbf{u}_y + A_z(x, t)\mathbf{u}_z$, the transverse components $A_y(x, t)$ and $A_z(x, t)$ must satisfy d'Alembert's equation $\frac{\partial^2 A_i}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 A_i}{\partial t^2} = 0$, with velocity $v = 1/\sqrt{\epsilon\mu}$, whose general solution is known to be

$$\underbrace{A_i(x, t)}_{\text{progressive wave}} = \underbrace{f_i \left(t - \frac{x}{v} \right)}_{\text{direction } \mathbf{u}_x} + \underbrace{g_i \left(t + \frac{x}{v} \right)}_{\text{direction } -\mathbf{u}_x}$$

where f_i and g_i are any functions. We are now looking for progressive plane waves propagating along \mathbf{u}_x toward positive x , that is we do not consider the second term $g_i \left(t + \frac{x}{v} \right)$ which can always be

²Static fields satisfying Maxwell's equations can always be added a posteriori to time-dependent solutions thanks to the linearity of Maxwell's equations and the superposition theorem.

added a posteriori thanks to the superposition theorem³. The electric field transverse components then read

$$\begin{aligned} A_y(x, t) &= f_y \left(t - \frac{x}{v} \right) \\ A_z(x, t) &= f_z \left(t - \frac{x}{v} \right) \end{aligned} \quad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} \quad \rightarrow \quad \begin{cases} E_y(x, t) = -f'_y \left(t - \frac{x}{v} \right) \\ E_z(x, t) = -f'_z \left(t - \frac{x}{v} \right) \end{cases}$$

The magnetic field components are now obtained as

$$\mathbf{B} = \nabla \times \mathbf{A} = \begin{vmatrix} \partial_x & 0 & A_y \\ 0 & \partial_x & A_z \\ 0 & A_y & \partial_x \end{vmatrix} \Rightarrow \begin{aligned} B_y(x, t) &= \frac{1}{v} f'_z(t - \frac{x}{v}) \\ B_z(x, t) &= -\frac{1}{v} f'_y(t - \frac{x}{v}) \end{aligned}$$

Now we expressed all the field transverse components and can check easily that $\mathbf{E} \cdot \mathbf{B} = 0$ (calculate $E_y B_y + E_z B_z$) and

$$\boxed{\mathbf{B} = \frac{\mathbf{u}_x}{v} \times \mathbf{E}},$$

as shown by the calculation $\frac{1}{v} \begin{vmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{v} f'_z & +f'_z \\ 0 & -f'_y & -f'_y \end{vmatrix} = \frac{1}{v} \begin{vmatrix} 0 & 0 & 0 \\ 0 & +f'_z & -f'_y \\ 0 & -f'_y & -f'_y \end{vmatrix}$. \mathbf{E} , \mathbf{B} , and \mathbf{u}_x thus form a direct trihedral and

$|\mathbf{B}| = |\mathbf{E}|/v$, which completes the derivation of the general properties of progressive plane waves propagating in linear matter.

10.6.6 Sinusoidal plane waves

Dealing with sinusoidal plane waves propagating in matter is not different from the case of vacuum. The complex notation greatly simplifies calculations with respect to using real fields, that require trigonometric manipulations. This section illustrates how to properly use complex notations in Maxwell's equation for describing sinusoidal plane waves.

Consider a sinusoidal plane wave, that is, a signal at any point of space \mathbf{x} and time t ,

$$\mathbf{s} = \mathbf{s}_0 \cos(\omega t - kx),$$

whose complex counterpart is

$$\underline{\mathbf{s}} = \underline{\mathbf{s}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad (10.42)$$

If \mathbf{s} has linear, circular or elliptic polarization, we can use a Cartesian basis to express the components of \mathbf{s} , which are then complex numbers:

$$\underline{\mathbf{s}}_0 = \underline{s}_{0x} \mathbf{u}_x + \underline{s}_{0y} \mathbf{u}_y + \underline{s}_{0z} \mathbf{u}_z$$

where $\underline{s}_{0x}, \underline{s}_{0y}, \underline{s}_{0z} \in \mathbb{C}$, and are independent of x, y, z or t .

³For instance, this necessary if there is a reflection on a mirror leading to a superposition of the forward and backward propagating waves.

Here we do not assume a particular direction for the wavevector and we chose it to be real, $\mathbf{k} = k_x \mathbf{u}_x + k_y \mathbf{u}_y + k_z \mathbf{u}_z$, $k_x, k_y, k_z \in \mathbb{R}$. This is not a constraint of the complex notation. Later, in order to properly describe absorption, we will choose $k_x, k_y, k_z \in \mathbb{C}$ without loss of generality.

How to apply the standard vectorial differentiation operators to the complex signal (10.42)? The answer is very simple due to the differentiation rule of the exponential function. The box below summarizes the effect of the usual vectorial differentiation operators on $\underline{\mathbf{s}}$

$$\begin{aligned}\frac{\partial \underline{\mathbf{s}}}{\partial t} &= -i\omega \underline{\mathbf{s}} & \nabla \cdot \underline{\mathbf{s}} &= i\mathbf{k} \cdot \underline{\mathbf{s}} \\ \frac{\partial \underline{\mathbf{s}}}{\partial x} &= ik_x \underline{\mathbf{s}} & \nabla \times \underline{\mathbf{s}} &= i\mathbf{k} \times \underline{\mathbf{s}} \\ && \Delta \underline{\mathbf{s}} &= -k^2 \underline{\mathbf{s}}\end{aligned}$$

To illustrate the simplicity of calculations with sinusoidal plane waves in complex notations, consider the propagation of a sinusoidal plane wave in a linear, homogeneous, isotropic and non-magnetic dielectric (ϵ, μ_0), without charge or current Electromagnetic waves are described by \mathbf{E}, \mathbf{B} satisfying Maxwell's equation

$$\begin{cases} \nabla \cdot \mathbf{B} = 0, \\ \nabla \cdot \mathbf{E} = 0, \end{cases} \quad \begin{cases} \nabla \times \mathbf{E} = -\partial_t \mathbf{B} \\ \nabla \times \mathbf{B} = v^{-2} \partial_t \mathbf{E}, \end{cases} \quad \text{with } \epsilon \mu_0 v^2 = 1$$

Looking for solutions in the form of sinusoidal plane waves amounts to applying the rules of vectorial differentiation to Maxwell's equations:

$$\begin{aligned}\underline{\mathbf{E}} &= \underline{\mathbf{E}}_0 e^{i\mathbf{k} \cdot \mathbf{r}} e^{-i\omega t} & \rightarrow & \begin{cases} i\mathbf{k} \cdot \underline{\mathbf{B}} = 0, \\ i\mathbf{k} \cdot \underline{\mathbf{E}} = 0, \end{cases} & \begin{cases} i\mathbf{k} \times \underline{\mathbf{E}} = i\omega \underline{\mathbf{B}} \\ i\mathbf{k} \times \underline{\mathbf{B}} = v^{-2} (-i\omega) \underline{\mathbf{E}} \end{cases} \\ \underline{\mathbf{B}} &= \underline{\mathbf{B}}_0 e^{i\mathbf{k} \cdot \mathbf{r}} e^{-i\omega t} \\ \mathbf{k} &= kn \quad \rightarrow \quad \begin{cases} \mathbf{B} \text{ is transverse} \\ \mathbf{E} \text{ is transverse} \end{cases} & \begin{cases} \mathbf{B} = (k/\omega) \mathbf{n} \times \mathbf{E} \\ \mathbf{E} = -v^2 (k/\omega) \mathbf{n} \times \mathbf{B} \end{cases}\end{aligned}$$

Writing $\mathbf{k} = kn$ to single out the unit vector \mathbf{n} along the propagation direction, we see that we obtain two relations showing that \mathbf{E} , \mathbf{B} and \mathbf{n} form a direct trihedral, but we do not know yet if they are equivalent to the general relation $\mathbf{B} = (1/v)\mathbf{n} \times \mathbf{E}$. Now, eliminating \mathbf{B} by combining these equations, we find

$$\mathbf{E} = -v^2 k^2 / \omega^2 \quad \underbrace{\mathbf{n} \times (\mathbf{n} \times \mathbf{E})}_{\text{Expand triple product}} \quad \rightarrow \quad \mathbf{E} = \frac{v^2 k^2}{\omega^2} \mathbf{E} \quad \rightarrow \quad \boxed{\omega = kv} \quad \text{dispersion relation}$$

In other words, we obtain the dispersion relation connecting ω and k , which can be finally introduced again into $\mathbf{B} = (k/\omega) \mathbf{n} \times \mathbf{E}$ to show that it coincides with the general relation.

The dispersion relation showing a linear relation between the frequency and the wavenumber of the wave, allows us to derive the phase velocity of the wave

$$v_\phi = \frac{\omega}{k} = v = \frac{1}{\sqrt{\epsilon \mu_0}} = \frac{c}{n}$$

where $n = \sqrt{\epsilon_r}$ is identified as the refractive index of the dielectric.

If the use of complex notation makes vectorial differentiation particularly simple for sinusoidal plane waves, care must be taken for nonlinear calculations. For example, the evaluation the Poynting vector

associated with a sinusoidal plane waves is worth commenting. There are two ways to calculate the Poynting vector properly:

Option 1: Keep real fields.

The Poynting vector is defined as

$$\boldsymbol{\Pi} = \mathbf{E} \times \mathbf{H}$$

A sinusoidal plane wave corresponds to the electromagnetic field

$$\mathbf{E} = \mathbf{E}_0 \cos(kx - \omega t), \quad \mathbf{H} = \mathbf{H}_0 \cos(kx - \omega t), \quad \text{where} \quad \mathbf{H}_0 = \frac{1}{\mu_0 v} \frac{\mathbf{k}}{k} \times \mathbf{E}_0$$

The Poynting vector thus reads

$$\boldsymbol{\Pi} = \mathbf{E}_0 \times \mathbf{H}_0 \cos^2(kx - \omega t) = \frac{E_0^2}{\mu_0 v} \frac{\mathbf{k}}{k} \cos^2(kx - \omega t).$$

As for the case of vacuum, at optical frequencies, the oscillation that appear in the Poynting vector are too fast to be seen by any detector, hence, we are interested in the time-average energy flux, defined as

$$\boldsymbol{\Pi}_{\text{ave}} = \langle \boldsymbol{\Pi}(t) \rangle = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} \boldsymbol{\Pi}(t) dt,$$

which is easy to evaluate if we remember that $\langle \cos^2(\cdot) \rangle = 1/2$:

$$\boldsymbol{\Pi}_{\text{ave}} = \frac{E_0^2}{2\mu_0 v} \frac{\mathbf{k}}{k}$$

$\boldsymbol{\Pi}_{\text{ave}}$ \equiv average em energy flux escaping the volume Ω through $\partial\Omega$.

Option 2: Complex Poynting vector.

The complex Poynting vector is defined by

$$\underline{\boldsymbol{\Pi}} = \frac{1}{2} \underline{\mathbf{E}} \times \underbrace{\underline{\mathbf{H}}^*}_{\text{c.c.}}$$

where the star denotes the complex conjugate and the coefficient $\frac{1}{2}$ is important as it takes care of the time average of the fast oscillating term ($\propto \cos^2(\omega t - kx)$). The complex Poynting vector is particularly easy to evaluate since all exponential factors cancel out (if we the wavevector components are purely real) by using the complex fields $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i\mathbf{k} \cdot \mathbf{x}} e^{-i\omega t}$, $\underline{\mathbf{H}} = \underline{\mathbf{H}}_0 e^{i\mathbf{k} \cdot \mathbf{x}} e^{-i\omega t}$, where $\underline{\mathbf{H}}_0 = \frac{1}{v\mu_0} \frac{\mathbf{k}}{k} \times \underline{\mathbf{E}}_0$:

$$\underline{\boldsymbol{\Pi}} = \frac{1}{2} \underline{\mathbf{E}} \times \underline{\mathbf{H}}^* = \frac{1}{2v\mu_0} \underline{\mathbf{E}} \times \left(\frac{\mathbf{k}}{k} \times \underline{\mathbf{E}}^* \right) = \frac{1}{2v\mu_0} (\mathbf{E} \cdot \mathbf{E}^*) \frac{\mathbf{k}}{k} = \frac{|\underline{\mathbf{E}}_0|^2}{2\mu_0 v} \frac{\mathbf{k}}{k}.$$

We find the same result as the time-average Poynting vector

$$\underline{\boldsymbol{\Pi}} = \langle \boldsymbol{\Pi}(t) \rangle,$$

which means that the time average electromagnetic power flux through a surface S is equal to the (real part of)⁴ the complex flux $\iint_S \underline{\boldsymbol{\Pi}} \cdot \mathbf{n} dS$.

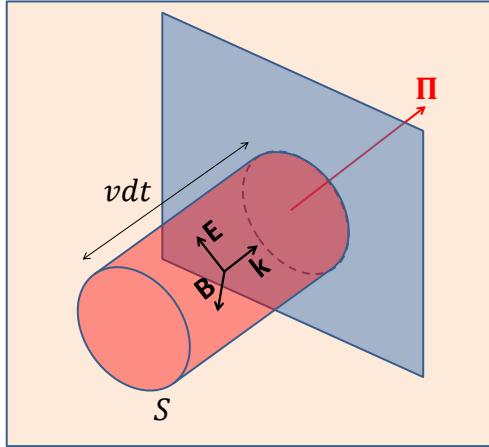
⁴Taking the real part will be necessary only in media with absorption, for which some components of the wavevector will be complex).

Now let's check that using the complex Poynting vector is compatible with what we know about energy conservation and calculate the time average density of electromagnetic energy that is supposed to cross a surface S during an infinitesimal time dt . Here again, the density of electromagnetic energy is a nonlinear function of the fields and a coefficient $1/2$ must be introduced in the complexified version of its expression, in order to take care of the time average procedure. The electric and magnetic energy densities then read

$$u_{e,\text{ave}} = \frac{1}{2}\epsilon \frac{\underline{\mathbf{E}} \cdot \underline{\mathbf{E}}^*}{2} = \frac{\epsilon |\underline{\mathbf{E}}_0|^2}{4}, \quad u_{m,\text{ave}} = \frac{1}{2}\mu_0 \frac{\underline{\mathbf{H}} \cdot \underline{\mathbf{H}}^*}{2} = \frac{1}{4v^2\mu_0} |\underline{\mathbf{E}}_0|^2$$

Remember that velocity is defined by $\epsilon\mu_0 v^2 = 1$. This allows us to calculate u_{ave} , the total time average density of electromagnetic energy:

$$u_{e,\text{ave}} = u_{m,\text{ave}} \quad \rightarrow \quad u_{\text{ave}} = \frac{\epsilon |\underline{\mathbf{E}}_0|^2}{2}.$$



Now the electromagnetic energy flowing through a surface S during dt , where the surface is perpendicular to the energy flow, is the energy contained in a volume $Svdt$, that is

$$dU_{\text{ave}} = u_{\text{ave}} S v dt = \frac{\epsilon v S |\underline{\mathbf{E}}_0|^2 dt}{2}$$

We see that it exactly corresponds to the flux of the complex Poynting vector through the surface

$$\Phi_{\underline{\Pi}} = \iint \underline{\Pi} \cdot \mathbf{n} dS = \frac{|\underline{\mathbf{E}}_0|^2}{2\mu_0 v} S = \frac{\epsilon v S}{2} |\underline{\mathbf{E}}_0|^2$$

To conclude on the conservation of electromagnetic energy

$$\boxed{\frac{dU_{\text{ave}}}{dt} = \Phi_{\underline{\Pi}}}.$$

- The complex Poynting vector represents for a sinusoidal plane wave the time-average current density of electromagnetic energy.
- Its flux through a surface corresponds to the time average electromagnetic power (energy during 1 second) that is transported by the sinusoidal plane wave through this surface.

10.6.7 Propagation in dielectrics

Since linear, homogeneous and isotropic media have a small magnetic susceptibility⁵, we can often restrict the study of wave propagation in matter to the case of nonmagnetic materials, for which $\mu_r = 1$. This is a very frequent case for most dielectric materials used for instance in modern laser technology⁶. In dielectrics, wave propagates at velocity

$$v = \frac{c}{\sqrt{\epsilon_r}} = \frac{c}{n},$$

where $n = \sqrt{\epsilon_r} > 1$ is the refractive index of the medium and $\epsilon_r = \epsilon/\epsilon_0$ the dielectric constant.

In addition to the general properties of progressive plane waves⁷, sinusoidal plane waves satisfy a *dispersion relation*, that is, a relation between the frequency ω and wavenumber k , that is found by introducing the expression for a plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(kx - \omega t)}$ into Eq. (10.37), leading to the following dispersion relation

$$k = n \frac{\omega}{c} = \frac{\omega}{v}.$$

It is almost the same as in vacuum and means that all frequencies propagate at the same velocity v in the medium. If we consider a superposition of two sinusoidal plane waves having different frequencies and propagating in the same direction, both waves will propagate at the same velocity. The crests of a wave will remain at fixed distance from the crests of the other during propagation.

As we will see, these results hold more generally when dealing with quasi-monochromatic wavepackets, that is, a superposition of waves that contain only a few frequency components, and far from any atomic resonance. The complete picture is, however, slightly more complicated: all charges (free and bound) have inertia and cannot adapt instantaneously to the variations of the oscillating electric field, so that $\mathbf{D}(\mathbf{x}, t) \neq \epsilon \mathbf{E}(\mathbf{x}, t)$. A natural consequence of this is that the permittivity ϵ depends on the frequency ω of the electromagnetic wave, leading to dispersion. This phenomenon explains for example the dispersion of sunlight by water droplets which is at the origin of rainbows. Dispersion will be discussed in detail in an appendix to the next chapter.

Additionally, a significant fraction of the electromagnetic energy may be absorbed if its frequency ω is close to the resonance frequencies of the atoms or molecules that constitute the medium, or more generally, if it is larger than the frequencies capable of inducing interband transitions in solids. We have all experienced this phenomenon: the semiconducting materials in our cellphones absorb visible light and this makes it possible for us to take pictures.

Absorption is actually simple to take into account when working with the complex representation of the fields. In the most general case, the response to a complex plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(kx - \omega t)}$ will lead

⁵Paramagnetic materials: $\chi_m \lesssim 10^{-3}$. Diamagnetic materials: $|\chi_m| \lesssim 10^{-5}$.

⁶Anisotropic materials are also frequently used

⁷The electromagnetic field for a plane wave is transverse, i.e., \mathbf{E} and \mathbf{B} are both perpendicular to the direction of propagation \mathbf{u}_x given by the wavevector $\mathbf{k} = k \mathbf{u}_x$, and

$$\mathbf{B} = \frac{\mathbf{k}}{\omega} \times \mathbf{E} = \frac{1}{v} \mathbf{u}_x \times \mathbf{E}.$$

to a frequency-dependent, complex permittivity (an explicit model for this frequency dependence will be developed in the next chapter

$$\underline{\epsilon}(\omega) = \epsilon_R(\omega) + i\epsilon_I(\omega),$$

so that in the absence of free charges and currents, the propagation equation (10.37) for a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\underline{k}x - \omega t)}$ gives the dispersion relation

$$\underline{k}^2 = \frac{\omega^2}{c^2} \frac{1}{\epsilon_0} (\epsilon_R(\omega) + i\epsilon_I(\omega)),$$

or, in terms of the complex refractive index $\underline{n} = n_R + in_I = \sqrt{\frac{\epsilon}{\epsilon_0}}$

$$\underline{k} = \frac{\omega}{c} (n_R(\omega) + in_I(\omega)) = k_R + ik_I,$$

where the imaginary part of the wavenumber characterizes the absorption of the electromagnetic wave by matter. Indeed, replacing \underline{k} in the complex plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\underline{k}x - \omega t)}$ and taking the real part, we have

$$\mathbf{E} = \mathbf{E}_0 \cos(k_R x - \omega t) e^{-k_I x} \quad (10.43)$$

and the amplitude of the wave decays exponentially with a typical decay length $\delta = \frac{1}{k_I}$. More details will be provided in the next chapter.

10.7 Summary and essential formulas

- Maxwell's equations form a complete set of coupled equations relating the electric and magnetic fields to their sources (charge ρ and current density \mathbf{J}):

$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$	Gauss-Maxwell,
$\nabla \cdot \mathbf{B} = 0$	Thomson-Maxwell,
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday-Maxwell,
$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$	Ampère-Maxwell,

which, together with the Lorentz force acting on a point charge q with velocity \mathbf{v} ,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

and an evolution equation for the electron velocity (for example Newton's equation), provide a complete description of all classical electromagnetism.

- In vacuum ($\rho = 0, \mathbf{J} = \mathbf{0}$), the electric and magnetic fields satisfy the same propagation equation

$$\begin{aligned}\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E} &= 0, \\ \nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{B} &= 0,\end{aligned}$$

where $c = 1/\sqrt{\epsilon_0 \mu_0} \approx 3 \times 10^8 \text{ m s}^{-1}$ is the speed of light. Each component $F(\mathbf{x}, t)$ of the electric and magnetic fields then satisfies d'Alembert's wave equation

$$\nabla^2 F(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 F(\mathbf{x}, t)}{\partial t^2} = 0.$$

- A plane wave propagating along the x -axis corresponds to a solution of d'Alembert's equation which depends solely on x and t . A preferred solution corresponds to a sinusoidal plane wave propagating toward positive values of x ,

$$f_+(x, t) = A_1 \cos(kx - \omega t + \phi_1),$$

where the spatial (k) and temporal (ω) frequencies are related by the dispersion relation

$$\omega(k) = ck.$$

- For a plane wave propagating along the direction \mathbf{n} in vacuum, the electric and magnetic fields are transverse:

$$\mathbf{E} \cdot \mathbf{n} = \mathbf{B} \cdot \mathbf{n} = 0, \tag{10.44}$$

and perpendicular to each other, so that $\{\mathbf{n}, \mathbf{E}, \mathbf{B}\}$ form a direct trihedral:

$$\mathbf{B} = \frac{1}{c} \mathbf{n} \times \mathbf{E}. \quad (10.45)$$

- A sinusoidal plane wave propagating along the x -axis accepts two independent solutions, linearly polarized along Oy and Oz , respectively

$$\begin{aligned} \mathbf{E}_y &= E_{0y} \cos(kx - \omega t + \phi_{0y}) \mathbf{u}_y, \\ \mathbf{E}_z &= E_{0z} \cos(kx - \omega t + \phi_{0z}) \mathbf{u}_z, \end{aligned}$$

and the most general polarization state is characterized by the set of parameters

$$\{E_{0y}, \phi_{0y}, E_{0z}, \phi_{0z}\}.$$

- The conservation equation for the total electromagnetic energy is given by Poynting's theorem

$$\iiint_{\Omega} \frac{\partial u_{EM}(\mathbf{x}, t)}{\partial t} d^3x = - \iiint_{\Omega} \mathbf{J} \cdot \mathbf{E} d^3x - \iint_{\partial\Omega} \mathbf{\Pi} \cdot d\mathbf{S}$$

or, equivalently, in its local form

$$\frac{\partial u_{EM}}{\partial t} + \nabla \cdot \mathbf{\Pi} = -\mathbf{J} \cdot \mathbf{E},$$

where the Poynting vector,

$$\mathbf{\Pi} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0},$$

gives the direction of the electromagnetic energy flow and

$$u_{EM} = \frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0}$$

is the density of electromagnetic energy.

- Due to the high frequency of optical waves, every detector of surface S oriented perpendicularly to the direction of propagation of the wave measures a time-averaged power given by $\mathcal{I}S$, where the intensity (also called irradiance) \mathcal{I} (in W m^{-2}) is

$$\mathcal{I} = \langle |\mathbf{\Pi}| \rangle_T.$$

For a sinusoidal plane wave of amplitude E_0 propagating along the x -axis, the time-averaged energy density and Poynting's vector are given by

$$\begin{aligned} \langle u_{EM} \rangle_T &= \frac{\epsilon_0 E_0^2}{2}, \\ \langle \mathbf{\Pi} \rangle_T &= c \langle u_{EM} \rangle_T \mathbf{u}_x = \mathcal{I} \mathbf{u}_x. \end{aligned}$$

- Maxwell's equations in matter may be rewritten as

$\nabla \cdot \mathbf{D} = \rho$	Gauss-Maxwell,
$\nabla \cdot \mathbf{B} = 0$	Thomson-Maxwell,
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	Faraday-Maxwell,
$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$	Ampère-Maxwell,

(10.46)

where the auxiliary fields are given by

$$\begin{aligned}\mathbf{D} &= \epsilon_0 \mathbf{E} + \mathbf{P}, \\ \mathbf{H} &= \frac{\mathbf{B}}{\mu_0} - \mathbf{M}.\end{aligned}$$

The behavior of the different types of materials is thus determined by the constitutive relations

$$\mathbf{P}(\mathbf{E}) \quad \text{and} \quad \mathbf{M}(\mathbf{B}).$$

Poynting's vector in matter reads, in a real representation,

$$\mathbf{\Pi} = \mathbf{E} \times \mathbf{H} \quad (10.47)$$

and the density of electromagnetic energy for a linear medium,

$$u_{EM} = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}).$$

- In a linear homogeneous and isotropic medium and in the absence of conduction charge or current, the scalar and vector potentials, the electric field and the magnetic field all satisfy the d'Alembert wave equation

$$\nabla^2 \mathbf{F} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \mathbf{F} = 0,$$

where $v = 1/\sqrt{\epsilon\mu}$ and ϵ and μ are the permittivity and permeability of the medium.

- Plane sinusoidal waves have the same structure in linear homogeneous and isotropic media except that they propagate at $v < c$ rather than c , the speed of light in vacuum. The electric and magnetic fields are transverse to the propagation direction \mathbf{n} and

$$\mathbf{B} = \frac{1}{v} \mathbf{n} \times \mathbf{E}.$$

The wavenumber $\mathbf{k} = kn$ is linked to the wave frequency ω by the dispersion relation

$$\omega = kv.$$

- The complex representation of the Poynting vector reads

$$\underline{\mathbf{\Pi}} = \frac{1}{2} \underline{\mathbf{E}} \times \underline{\mathbf{H}}^*.$$

Chapter 11

Electromagnetic waves in metals and dielectrics, waveguides

Introduction

This chapter contains two intertwined parts which describe the propagation of electromagnetic waves in metals on the one hand and in dielectrics on the other hand. Maxwell's equations allow us to describe both cases with a common framework. However, each type of medium satisfies a specific equation of state, leading to specific properties for the propagation of electromagnetic waves. For instance the dispersion relation will depend on the equation of state of the medium. As we will see, there exist a strong similarity between wave propagation in these media, which is due to the fact that the distinction between free and bound charges that was made from their response to permanent fields becomes less obvious for oscillating fields.

Both cases have a common application: the propagation of electromagnetic waves in waveguides. Metallic waveguides are typically used for the propagation of micro-waves of frequency between 300 MHz (wavelength of 1 m) and 300 GHz (wavelength of 1 mm) but the physical mechanism underlying the guiding phenomenon is also valid for radio-waves with frequencies down to 30 Hz. For instance, radio waves are guided between an emitter and a receiver by the surface of the Earth and the ionosphere where ions in the ionosphere act as a metallic medium reflecting the electromagnetic waves, thereby ensuring waveguiding. Light and more generally electromagnetic waves of high frequency can be guided by optical fibers, which work for electromagnetic waves of high frequency similarly to metallic waveguides for microwaves, that is, by internal reflection of the waves on the wall of the fiber, allowing for a zig-zag propagation of electromagnetic waves in the fiber.

This chapter presents general properties of electromagnetic waves propagating in conductors and in dielectrics. Properties allowing for waveguiding by metallic conductors and dielectrics are reviewed. The physical origin of absorption and its link with dispersion is discussed. In an appendix section, as a transition to the Optics course, we will dwell more upon dispersion and show how light pulses in matter can be stretched or compressed due to the phenomenon of dispersion which is inherent to the inertia of charged particles.

11.1 Propagation of waves in media with absorption

11.1.1 Summary on wave propagation equations and their plane wave solutions

In chapter 10, we established the propagation equations for the electric and the magnetic fields in the general case of a linear, homogeneous and isotropic medium of permittivity ϵ and permeability μ , which can include conduction charges of density $\rho(\mathbf{r}, t)$ and currents of density $\mathbf{J}(\mathbf{r}, t)$. We found for the electric field:

$$\nabla^2 \mathbf{E} - \epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\nabla\rho}{\epsilon} + \mu \frac{\partial \mathbf{J}}{\partial t}. \quad (11.1)$$

and for the magnetic field:

$$\nabla^2 \mathbf{B} - \epsilon\mu \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu \nabla \times \mathbf{J}, \quad (11.2)$$

In the absence of right-hand side, that is, in media without conduction charges or currents, equations (11.1) and (11.2) become the homogeneous D'alembert wave equations, from which we inferred that the structure of electromagnetic plane waves in LHI media is similar to that in vacuum, simply replacing light velocity in vacuum by its value in the medium $v = 1/\sqrt{\epsilon\mu}$. In particular, \mathbf{E} , \mathbf{B} and the propagation direction \mathbf{n} form a direct trihedron and $\mathbf{B} = v^{-1}\mathbf{n} \times \mathbf{E}$.

We then showed that a plane wave $\underline{\mathbf{E}} = \mathbf{E}_0 e^{i(kx-\omega t)}$ is a solution to Eq. (11.1) (without right-hand side) provided the dispersion relation

$$k = \frac{\omega}{v}$$

is satisfied, where both the frequency and the wavenumber k are real valued, which describes a situation where all frequencies propagate at the same velocity v .

At the end of chapter 10, we have very briefly discussed the fact that dielectric media do not instantaneously respond to the variations of the oscillating electric field. The inertia of charges induces a delay between the excitation (the electric field) and the response (the polarization). The consequence is that the generalization of the equation of state $\mathbf{D}(\mathbf{r}, t) = \epsilon \mathbf{E}(\mathbf{r}, t)$ for a LHI dielectric is a relation where the permittivity depends on the frequency of the wave¹

$$\underline{\mathbf{D}}(\mathbf{r}, \omega) = \underline{\epsilon}(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega),$$

where the fields $\mathbf{D}(\mathbf{r}, t)$ and $\mathbf{E}(\mathbf{r}, t)$ represent the complex amplitudes of a sinusoidal plane wave of frequency ω . The permittivity $\underline{\epsilon}(\omega)$ is not only frequency dependent but also complex valued. Physically, this means the the wave will undergo both dispersion and absorption during propagation. An explicit model for this frequency dependence of the complex permittivity was developed in lecture 5 and is reminded in section 11.3.1 of the present chapter, while the phenomenon of dispersion is discussed in more detail in the appendix section 11.10.

¹It is in fact the same equation of state expressed in the frequency (Fourier) domain, where time-dependent fields are Fourier transformed into frequency-dependent fields.

Similarly, in magnetized LHI media, the equation of state becomes

$$\underline{\mathbf{B}}(\mathbf{r}, \omega) = \underline{\mu}(\omega) \underline{\mathbf{H}}(\mathbf{r}, \omega),$$

where $\underline{\mu}(\omega)$ denotes the frequency-dependent complex permeability of the medium, and for metals, we can generalize Ohm's law to take into account the inertia of charges as

$$\underline{\mathbf{J}}(\mathbf{r}, \omega) = \underline{\gamma}(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega),$$

where $\underline{\gamma}(\omega)$ denotes the frequency-dependent complex conductivity. The Drude-Lorentz model presented in section 11.2.2 constitutes an explicit model for the latter quantity, which also gives rise to dispersion and absorption. For metals, absorption is in any case obtained even for a constant conductivity.

We will focus in this chapter on non magnetized media, and in particular on the case of metals carrying only conduction currents and of dielectrics. Both cases lead to absorption, resulting in the attenuation of waves along propagation. For a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$ for instance in a dielectric without free charges or currents, the propagation equation (11.1) can be rewritten in complex notations by using the complex permittivity $\underline{\epsilon}(\omega) = \epsilon_R(\omega) + i\epsilon_I(\omega)$, leading to

$$\nabla^2 \underline{\mathbf{E}} - \mu_0 \underline{\epsilon}(\omega) \frac{\partial^2 \underline{\mathbf{E}}}{\partial t^2} = \nabla^2 \underline{\mathbf{E}} - \frac{\underline{n}^2(\omega)}{c^2} \frac{\partial^2 \underline{\mathbf{E}}}{\partial t^2} = 0.$$

Writing $\mathbf{k} = \underline{k}\mathbf{n}$, with \mathbf{n} the direction of propagation, yields the dispersion relation

$$\underline{k}^2 = \omega^2 \mu_0 (\epsilon_R(\omega) + i\epsilon_I(\omega)).$$

showing that the wavenumber is complex valued as well. Defining a complex valued refractive index as

$$\underline{n}^2(\omega) = \frac{\underline{\epsilon}(\omega)}{\epsilon_0} = \frac{\epsilon_R(\omega) + i\epsilon_I(\omega)}{\epsilon_0},$$

the dispersion relation becomes

$$\underline{k} = \underline{n}(\omega) \frac{\omega}{c} = k_R + ik_I, \quad (11.3)$$

with $k_R(\omega) = n_R(\omega) \frac{\omega}{c}$ and $k_I = n_I(\omega) \frac{\omega}{c} > 0$.

As briefly discussed at the end of chapter 10, the imaginary part of the wavenumber $\underline{k} = k_R + ik_I$ characterizes absorption of the electromagnetic wave by matter. Indeed, replacing \underline{k} in the complex plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\underline{k}x - \omega t)}$, we have, assuming propagation along the x -axis

$$\underline{\mathbf{E}}(x, t) = \underline{\mathbf{E}}_0 \underbrace{e^{i(k_R x - \omega t)}}_{\text{plane wave, speed } c/n_R} \underbrace{e^{-k_I x}}_{\text{attenuation due to absorption}}. \quad (11.4)$$

and taking the real part,

$$\mathbf{E} = \mathbf{E}_0 \cos(k_R x - \omega t) e^{-k_I x}. \quad (11.5)$$

The amplitude of the wave decays exponentially with a typical decay length $\delta = \frac{1}{k_I}$.

11.1.2 Propagation with absorption

Once the wavenumber is allowed to be complex valued and as long as the equation of state remains linear, we see that there is only little change in the general properties of plane waves established in media without absorption: The structure of the wave is similar. The amplitude decays along propagation and special attention must be paid to the energy conservation equation and Poynting vector.

The magnetic field associated with the electric field (11.5) can be obtained in its complex representation with the help of the general relation for plane waves $\underline{\mathbf{B}} = \frac{\underline{\mathbf{k}}}{\omega} \times \underline{\mathbf{E}}$,

$$\underline{\mathbf{B}} = \frac{k}{\omega} \mathbf{u}_x \times \underline{\mathbf{E}}_0 e^{i(k_R x - \omega t)} e^{-k_I x}.$$

Writing $k = |k|e^{-i\phi}$, the real magnetic field reads

$$\mathbf{B} = \frac{|k|}{\omega} \cos(k_R x - \omega t - \phi) e^{-k_I x} (\mathbf{u}_x \times \mathbf{E}_0),$$

so that both fields do not oscillate in phase in the presence of absorption ($\phi \neq 0$ if $k_I \neq 0$). The Poynting vector is given by

$$\underline{\mathbf{\Pi}} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} = \frac{|\mathbf{E}_0|^2 |k|}{\mu_0 \omega} \underbrace{\cos(k_R x - \omega t - \phi) \cos(k_R x - \omega t)}_{\frac{1}{2}\{\cos[2(k_R x - \omega t) - \phi] + \cos \phi\}} e^{-2k_I x} \mathbf{u}_x,$$

whose time-average reads

$$\langle \underline{\mathbf{\Pi}} \rangle_T = \frac{E_0^2}{2\mu_0 \omega} \underbrace{|k| \cos \phi}_{k_R} e^{-2k_I x} \mathbf{u}_x = \underbrace{\frac{|\mathbf{E}_0|^2 k_R}{2\mu_0 \omega} e^{-2k_I x}}_{\mathcal{I}} \mathbf{u}_x.$$

Note that the same result may be easily obtained by using the complex Poynting vector $\underline{\mathbf{\Pi}} = \frac{\underline{\mathbf{E}} \times \underline{\mathbf{B}}^*}{2\mu_0}$.

Indeed,

$$\underline{\mathbf{\Pi}} = \frac{\underline{\mathbf{E}} \times \underline{\mathbf{B}}^*}{2\mu_0} = \frac{1}{2\mu_0 \omega} \underline{\mathbf{E}} \times (k^* \mathbf{u}_x \times \underline{\mathbf{E}}^*) = \frac{k^*}{2\mu_0 \omega} (\underline{\mathbf{E}} \cdot \underline{\mathbf{E}}^*) \mathbf{u}_x = \frac{k^* |\mathbf{E}_0|^2}{2\mu_0 \omega} e^{-2k_I x} \mathbf{u}_x$$

and the result

$$\langle \underline{\mathbf{\Pi}} \rangle_T = \text{Re}\{\underline{\mathbf{\Pi}}\} = \frac{k_R |\mathbf{E}_0|^2}{2\mu_0 \omega} e^{-2k_I x} \mathbf{u}_x.$$

still holds for a plane wave in an absorptive medium.

In a dielectric medium, the intensity of the wave, which is proportional to $\langle |\mathbf{E}|^2 \rangle$, therefore decays as

$$\mathcal{I}(x) = \frac{\epsilon_0 |\mathbf{E}_0|^2 n_r c}{2} e^{-2k_I x} = I_0 e^{-\alpha x},$$

where $n_r = \text{Re}(\sqrt{\epsilon(\omega)/\epsilon_0})$ and $\alpha = 2k_I$ (m^{-1}) is called the absorption coefficient. The fraction T of light transmitted through a medium of thickness d therefore follows Beer-Lambert's law,

$$T = e^{-\alpha d}.$$

In a metal with finite conductivity, a similar law applies, as detailed in the next section.

11.2 Electromagnetic waves in conductors

Consider the situation of an electromagnetic wave interacting with a homogeneous medium in the presence of free charges. For instance, the medium is a metal or a plasma (partially ionized gas). In both cases, a density $n \neq 0$ of free electrons is present and a current density \mathbf{J} will appear in response to the electric field of the wave, according to Ohm's law. Assuming that the medium preserves its neutrality², $\rho = 0$, and the propagation equation (10.9) for the electric field in a conductor becomes

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E} = \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (11.6)$$

If we assume a sinusoidal plane electromagnetic wave of the form

$$\mathbf{E}(x, t) = \operatorname{Re} \left\{ \underline{\mathbf{E}}_0 e^{i(\underline{k}x - \omega t)} \right\} = \operatorname{Re} \{ \underline{\mathbf{E}} \},$$

the current density in its complex representation is given by Ohm's law,

$$\underline{\mathbf{J}} = \underline{\sigma}(\omega) \underline{\mathbf{E}},$$

where the conductivity $\underline{\sigma}$ is a function of the wave frequency and in the general case, a complex number. Injecting $\underline{\mathbf{E}}_0 e^{i(\underline{k}x - \omega t)}$ into Eq. (11.6) gives

$$\left(-\underline{k}^2 + \frac{\omega^2}{c^2} \right) \underline{\mathbf{E}} = -\mu_0 \underline{\sigma}(\omega) i \omega \underline{\mathbf{E}}$$

and we obtain the following dispersion relation

$$\underline{k}^2 = \frac{\omega^2}{c^2} + i \omega \mu_0 \underline{\sigma}(\omega).$$

(11.7)

Again, this dispersion relation leads to a complex wavenumber $\underline{k} = k_R + ik_I$, so that

$$\underline{\mathbf{E}}(x, t) = \underline{\mathbf{E}}_0 e^{i(k_R x - \omega t)} e^{-ik_I x}.$$

In conclusion, waves in metals and dielectrics take the same general form. From this, we see that it is possible to characterize a conductor by an effective permittivity $\underline{\epsilon}$, or equivalently, a complex refractive index $\underline{n} = \sqrt{\underline{\epsilon}/\epsilon_0}$ such that $\underline{k} = \underline{n} \frac{\omega}{c}$ and so,

$$\underline{\epsilon}(\omega) = \epsilon_0 \underline{n}^2(\omega) = \epsilon_0 \left(1 + i \frac{\sigma(\omega)}{\epsilon_0 \omega} \right) = \epsilon_R(\omega) + i \epsilon_I(\omega)$$

Conversely, in a dielectric an oscillating polarization vector \mathbf{P} gives rise to a current density

$$\mathbf{J}_P = \frac{\partial \mathbf{P}}{\partial t},$$

²That is, the density n of positive ions is the same as that of electrons and the particle distribution is homogeneous.

which in complex notations reads

$$\underline{\mathbf{J}}_P = -i\omega \underline{\mathbf{P}} == -i\omega(\underline{\epsilon}(\omega) - \epsilon_0) \underline{\mathbf{E}}$$

and so, an effective conductivity can be defined as

$$\underline{\sigma}(\omega) = -i\omega(\underline{\epsilon}(\omega) - \epsilon_0).$$

The fact that we can relate a conductivity to a permittivity in the time-dependent regime is not surprising since, in reality, the distinction between bound and free electrons is somehow arbitrary in the presence of a sinusoidal electromagnetic wave. Both free and bound charges will oscillate around their mean positions generating an oscillating current in both cases. Calculating a conductivity $\underline{\sigma}$ or a permittivity $\underline{\epsilon}$ is therefore only a matter of convention. This, of course, does not mean that metals and dielectrics behave the same way. A main difference between them comes from the frequency-dependence of the permittivity/conductivity.

11.2.1 Losses in conductors

According to Poynting's theorem, an electromagnetic wave losses its energy due to the power dissipated in the medium. For a sinusoidal wave of frequency ω , the mean power dissipated per unit volume reads

$$\langle \mathcal{P} \rangle_T = \frac{1}{2} \text{Re}\{\underline{\mathbf{J}} \cdot \underline{\mathbf{E}}^*\}.$$

Using the complex conductivity $\underline{\sigma}$, we find

$$\langle \mathcal{P} \rangle_T = \frac{1}{2} \text{Re}\{\underline{\sigma}(\omega)\} |\underline{\mathbf{E}}|^2.$$

If instead we use the effective permittivity, the complex current density reads $\underline{\mathbf{J}} = -i\omega(\underline{\epsilon}(\omega) - \epsilon_0) \underline{\mathbf{E}}$ and we find

$$\langle \mathcal{P} \rangle_T = \frac{\omega}{2} \text{Re}\{-i(\underline{\epsilon}(\omega) - \epsilon_0)\} |\underline{\mathbf{E}}|^2 = \frac{\omega}{2} \text{Im}\{\underline{\epsilon}(\omega)\} |\underline{\mathbf{E}}|^2. \quad (11.8)$$

11.2.2 Conductivity in a time-dependent regime: the Drude-Lorentz model

Now we will derive the equation of motion of an electron of charge $-e$, mass m_e , oscillating around the position x_0 (we assume that the amplitude of the oscillations is negligible with respect to the wavelength of the wave, and so we treat the fields as spatially constants around x_0). Newton's equation for the electron reads

$$m_e \frac{d\mathbf{v}(t)}{dt} = -e (\mathbf{E}(x_0, t) + \mathbf{v} \times \mathbf{B}(x_0, t)) - \frac{m_e}{\tau} \mathbf{v},$$

where \mathbf{v} is the velocity of the electron and τ is the time between collisions. Assuming that the speed of the electron is much smaller than c we can neglect the magnetic component of the Lorentz force. Then, writing $\mathbf{v}(t) = \operatorname{Re} \{\underline{\mathbf{v}}_0 e^{-i\omega t}\} = \operatorname{Re} \{\underline{\mathbf{v}}\}$, we get

$$-i\omega m_e \underline{\mathbf{v}} = -e \underline{\mathbf{E}} - \frac{m_e}{\tau} \underline{\mathbf{v}}$$

and so,

$$\underline{\mathbf{v}} = \underline{\mathbf{v}}_0 e^{-i\omega t} = \frac{e/m_e}{i\omega - 1/\tau} \underbrace{\underline{\mathbf{E}}_0 e^{i(kx - \omega t)}}_{\underline{\mathbf{E}}}.$$

Finally, the current density associated to this oscillatory motion of the electrons is

$$\underline{\mathbf{J}} = -ne \underline{\mathbf{v}} = \frac{ne^2/m_e}{1/\tau - i\omega} \underline{\mathbf{E}},$$

where we have neglected the contribution of the ions, since they are several thousand times heavier than the electrons. Note that from this, the frequency-dependent conductivity reads

$$\underline{\sigma}(\omega) = \frac{ne^2\tau/m_e}{1 - i\omega\tau} = \frac{\sigma_0}{1 - i\omega\tau}, \quad (11.9)$$

where σ_0 is the usual static conductivity. Separating real and imaginary parts and replacing $\underline{\sigma}$ in the dispersion relation (11.7) yields

$$\underline{k}^2 = \frac{\omega^2}{c^2} \left(1 - \frac{\omega_p^2 \tau^2}{1 + \omega^2 \tau^2} + i \frac{\omega_p^2 \tau}{\omega(1 + \omega^2 \tau^2)} \right), \quad (11.10)$$

where the quantity ω_p is called the *plasma frequency* and is defined as

$$\omega_p = \sqrt{\frac{ne^2}{m_e \epsilon_0}}. \quad (11.11)$$

From Eq. (11.10), we see that it is possible to identify a complex permittivity for a metal (or a plasma), or equivalently a complex refractive index.

11.2.3 The effective refractive index of metals

The refractive index of a conductor then reads

$$\underline{n}^2(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\omega/\tau} = \frac{\epsilon(\omega)}{\epsilon_0} \quad (11.12)$$

In a metal, one has typically $1/\tau \sim 1 \times 10^{14} \text{ s}^{-1}$ so that in the visible range $\omega^2 \gg \omega/\tau$ and the refractive index becomes

$$\underline{n}^2(\omega) \approx 1 - \frac{\omega_p^2}{\omega^2}.$$

Since the plasma frequency of metals is typically in the UV region, in the visible range $\omega < \omega_p$, which implies that metals have a purely imaginary refractive index

$$\underline{n}(\omega) \approx i\sqrt{\frac{\omega_p}{\omega} - 1},$$

and the complex wavenumber $\underline{k} = \underline{n}\omega/c$ is purely imaginary as well, so visible light waves do not propagate at all in metals. These waves are completely reflected by metals and this is how metallic mirrors work. In reality, in real metals we should also include the contribution from bound electrons, so that the frequency dependent refractive index should read

$$\underline{n}^2(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\omega/\tau} + \omega_p'^2 \sum_k \frac{f_k}{\omega_{0k}^2 - \omega^2 - i\gamma_k\omega},$$

where the origin of the last term will become clear from the presentation of the Lorentz model in section 11.3.1. For now, simply note that for some metals, some of the resonance frequencies ω_{0k} may fall in the visible part of the spectrum (blue in the case of gold, red for copper), which then give them a particular color. If no absorption peaks are present in the visible range, a metal will have simply a silver color.

- **LOW FREQUENCY REGIME ($\omega \ll 1/\tau$)**

Consider now the low frequency regime, that is, $\omega\tau \ll 1$. The dynamics of the electrons is thus dominated by collisions, and the dispersion relation (11.10) becomes

$$\underline{k}^2 \approx \frac{\omega^2}{c^2} \left(1 - \omega_p^2 \tau^2 + i \frac{\omega_p^2 \tau}{\omega} \right).$$

Two distinct cases arise. If $\omega_p^2 \tau^2 \ll 1$, which is verified for a poor conductor, we can further write

$$\underline{k}^2 \approx \frac{\omega^2}{c^2} \left(1 + i \frac{\omega_p^2 \tau}{\omega} \right) = \frac{\omega^2}{c^2} \left(1 + i \frac{\sigma_0}{\epsilon_0 \omega} \right).$$

By writing the wave number as $k = k_R + ik_I$, we have $k^2 = k_R^2 - k_I^2 + 2ik_Rk_I$. Solving for k_R and k_I gives

$$\underline{k} = \frac{\omega}{c} \left(\sqrt{\frac{\sqrt{\epsilon_0^2 \omega^2 + \sigma_0^2} + \epsilon_0 \omega}{2\epsilon_0 \omega}} + i \sqrt{\frac{\sqrt{\epsilon_0^2 \omega^2 + \sigma_0^2} - \epsilon_0 \omega}{2\epsilon_0 \omega}} \right).$$

The skin depth then reads

$$\delta = \frac{1}{k_I} = \frac{c}{\omega} \sqrt{\frac{2\epsilon_0 \omega}{\sqrt{\epsilon_0^2 \omega^2 + \sigma_0^2} - \epsilon_0 \omega}},$$

which in the limit $\sigma_0^2 \gg \epsilon_0^2 \omega^2$ becomes

$$\delta \approx \sqrt{\frac{2\epsilon_0 c^2}{\sigma_0 \omega}} = \sqrt{\frac{2}{\sigma_0 \mu_0 \omega}}.$$

In the other case where $\omega_p^2\tau^2 \gg 1$, which is typically the case in metals, the dispersion relation (11.10) is approximated as

$$\underline{k}^2 \approx \frac{\omega^2}{c^2} \left(-\omega_p^2\tau^2 + i\frac{\omega_p^2\tau}{\omega} \right) \approx i\frac{\omega^2}{c^2} \frac{\omega_p^2\tau}{\omega}$$

and so,

$$\underline{k} = \frac{\omega}{c} \sqrt{\frac{\omega_p^2\tau}{\omega}} \left(\frac{1+i}{\sqrt{2}} \right).$$

The skin depth then reads

$$\delta = \frac{1}{k_I} = \sqrt{\frac{2c^2}{\omega_p^2\omega\tau}}.$$

Recalling that the static conductivity of a metal reads $\sigma_0 = \epsilon_0\omega_p^2\tau$, one can rewrite the skin depth as

$$\delta = \sqrt{\frac{2}{\mu_0\omega\sigma_0}}.$$

We recover the same low-frequency limit for the skin depth as in the case $\omega_p^2\tau^2 \ll 1$. In the following we consider the case of intermediate and high frequencies for metals, that is, we assume that we have $\omega\tau \gg 1$.

- **INTERMEDIATE FREQUENCY REGIME ($1/\tau \ll \omega \ll \omega_p$)**

For intermediate frequencies such that $1/\tau \ll \omega \ll \omega_p$ (in the case of a good conductor, this corresponds to $\omega \sim 10^{15} \text{ s}^{-1}$, which falls in the visible range), the imaginary part in the dispersion relation (11.10) can be neglected, so that

$$\underline{k}^2 = \frac{\omega^2}{c^2} \left(1 - \frac{\omega_p^2}{\omega^2} \right).$$

For frequencies below the plasma frequency ($\omega < \omega_p$) we obtain a pure imaginary wave number,

$$\underline{k} = i\frac{1}{c} \sqrt{\omega_p^2 - \omega^2}.$$

An incident electromagnetic wave is thus completely reflected at the interface and the skin depth is given by

$$\delta = \frac{1}{k_I} = \frac{c}{\sqrt{\omega_p^2 - \omega^2}}$$

- **HIGH FREQUENCY REGIME ($1/\tau \ll \omega_p \ll \omega$)**

Finally, in the high frequency regime where $\omega > \omega_p$, one obtains a real wave vector

$$k \approx \frac{\omega}{c} \left(\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \right).$$

The medium allows the propagation of electromagnetic waves without attenuation and the refractive index becomes an affine function of the electron density n_e , identified by the small- ω_p/ω expansion of k ,

$$k \approx \frac{\omega}{c} \left(1 - \frac{1}{2} \frac{\omega_p^2}{\omega^2} \right) = \underbrace{\frac{\omega}{c} \left(1 - \frac{1}{2} \frac{e^2}{m_e \epsilon_0 \omega^2} n_e \right)}_n.$$

Example 11.1 - Electromagnetic waves in copper and in the ionosphere

Consider the case of copper, whose free electron density is $n^{\text{Cu}} = 8.5 \times 10^{28} \text{ m}^{-3}$ and for which the mean time between collisions is $\tau^{\text{Cu}} \sim 10^{-14} \text{ s}$.

- a. What is the skin depth for an incident electromagnetic wave in the visible range ($\omega \sim 4 \times 10^{15} \text{ s}^{-1}$) and for a radio wave ($\omega \sim 6 \times 10^6 \text{ s}^{-1}$)?
- b. Compare with the case of the ionosphere, which is a plasma of low density $n^{\text{ion}} \sim 10^{12} \text{ m}^{-3}$ with conductivity $\sigma_0^{\text{ion}} \sim 1 \text{ S m}^{-1}$.

Solution

- a. We evaluate the plasma frequency $\omega_p = \sqrt{\frac{ne^2}{m_e \epsilon_0}}$ for copper

$$\omega_p^{\text{Cu}} = 1.6 \times 10^{16} \text{ s}^{-1}$$

and so, $\omega_p \tau \gg 1$. For copper in the visible range, $\omega \tau^{\text{Cu}} \sim 40$ and $\omega < \omega_p^{\text{Cu}}$, which corresponds to the intermediate frequency regime. The wave number is purely imaginary and the skin depth reads

$$\frac{1}{k_I} = \frac{c}{\sqrt{\omega_p^2 - \omega^2}} = 18 \text{ nm}.$$

For a radio wave, $\omega \tau^{\text{Cu}} \ll 1$ and therefore

$$\delta = \sqrt{\frac{2}{\mu_0 \omega \sigma_0}} = \sqrt{\frac{2c^2}{\omega_p^2 \omega \tau}} \sim 10 \text{ km}.$$

- b. For the ionosphere, the plasma frequency is

$$\omega_p^{\text{ion}} = 5.6 \times 10^7 \text{ s}^{-1},$$

whereas the time between collisions in the atmosphere can be obtained from the static conductivity, since

$$\sigma_0 = \frac{ne^2 \tau}{m_e} = \epsilon_0 \omega_p^2 \tau$$

and so,

$$\tau^{\text{ion}} \sim 3.5 \times 10^{-5} \text{ s}.$$

Once again $\omega_p \tau \gg 1$. The radio wave does not propagate in the ionosphere, since $\omega < \omega_p^{\text{ion}}$, and since $\omega \tau^{\text{ion}} \gg 1$

$$\delta = \frac{1}{k_I} = \frac{c}{\sqrt{\omega_p^2 - \omega^2}} = 5.3 \text{ m},$$

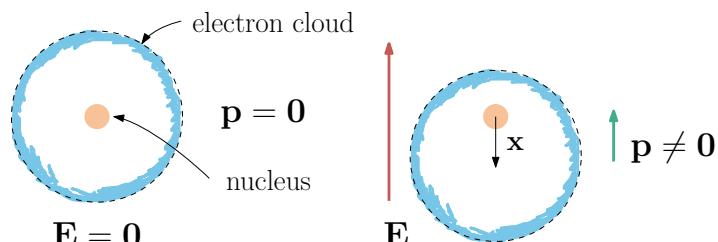
whereas visible light is such that $\omega \gg \omega_p$ and the ionosphere becomes transparent.

11.3 Electromagnetic waves in dielectric media

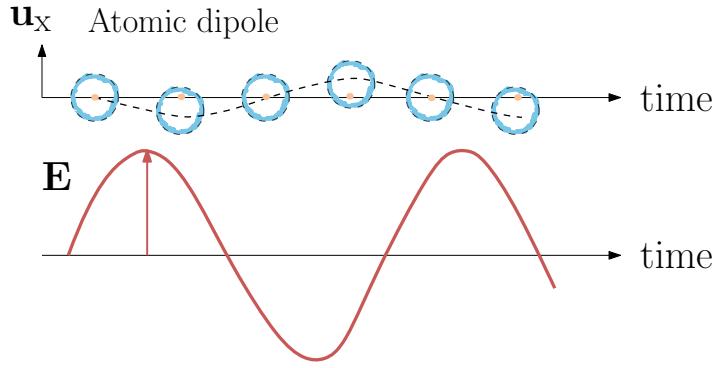
11.3.1 The Lorentz model for the classical atom (1878)

We will now establish a model for the frequency-dependence of the dielectric constant. This section is a reminder of lecture 5 and can be skipped in a first reading.

Let us recall that in an atom, under the presence of an electric field, the electron cloud orbiting around the nucleus will be shifted from its equilibrium position by a quantity \mathbf{x} . In consequence an atomic electric dipole \mathbf{p} develops. Here, we neglect the movement of the nucleus in the electric field due to its much larger mass compared to that of an electron.



In the presence of an electromagnetic wave, the electric field oscillates periodically, and therefore, so does the induced dipole. We expect that if the frequency ω of the wave is very low, the dipole will instantaneously adapt to the electric field and oscillate in phase with the latter, such that $\lim_{\omega \rightarrow 0} \epsilon(\omega) = \epsilon$, where ϵ is the static permittivity. In contrast, if the frequency is too high, the inertia of the electron cloud will make it impossible for the electrons to follow the rapid oscillations of the electric field, and the medium will become transparent. We then expect the permittivity to tend asymptotically to ϵ_0 as $\omega \rightarrow \infty$, that is, $\lim_{\omega \rightarrow \infty} \epsilon(\omega) = \epsilon_0$ and waves of very high frequency will propagate as if they were in vacuum.



The Lorentz model takes into account two crucial ingredients: firstly, since an oscillating dipole generates in fact an electric current, it creates its own electromagnetic wave. This means that a dipole driven at a frequency ω will lose part of its mechanical energy by emitting radiation at a frequency close to ω . Secondly, we know from experiments that electrons in atoms have discrete energy levels and as such, they absorb or emit light at precise resonant frequencies, a fact that arises from the quantum behavior of matter at atomic scales. The Lorentz model is a classical model that incorporates these two effects: the movement of an electron around the nucleus is modeled by a damped harmonic oscillator with a natural frequency ω_0 . The equation of motion reads

$$m_e \frac{d^2 \mathbf{x}}{dt^2} = -e \mathbf{E} - m_e \gamma \frac{d\mathbf{x}}{dt} - m_e \omega_0^2 \mathbf{x}$$

where \mathbf{x} is the displacement of the electron with respect to the equilibrium position around the nucleus and γ is a damping factor that accounts for the fact that energy is absorbed from the incident electromagnetic wave and re-emitted by the electrons (in an atom γ is typically of the order of 10^8 s^{-1}). Here we neglected the magnetic component of the Lorentz force since, for a plane electromagnetic wave and for velocities much smaller than the speed of light,

$$| -e\mathbf{v} \times \mathbf{B} | \leq e|\mathbf{v}| |\mathbf{B}| = e \frac{|\mathbf{v}|}{c} |\mathbf{E}| \ll e|\mathbf{E}|$$

Assuming now a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$ with a wavelength much longer than the oscillation amplitude of the atomic dipole, we may consider the field to be spatially uniform at the position of the dipole and look for a stationary solution in the form of a forced oscillator $\underline{\mathbf{x}} = \underline{\mathbf{x}}_0 e^{-i\omega t}$, so we have

$$-m_e \omega^2 \underline{\mathbf{x}} = -e \underline{\mathbf{E}} + i\omega m_e \gamma \underline{\mathbf{x}} - m_e \omega_0^2 \underline{\mathbf{x}}.$$

Then,

$$\underline{\mathbf{x}} = \frac{e}{m_e \omega^2 - \omega_0^2 + i\gamma\omega} \underline{\mathbf{E}}. \quad (11.13)$$

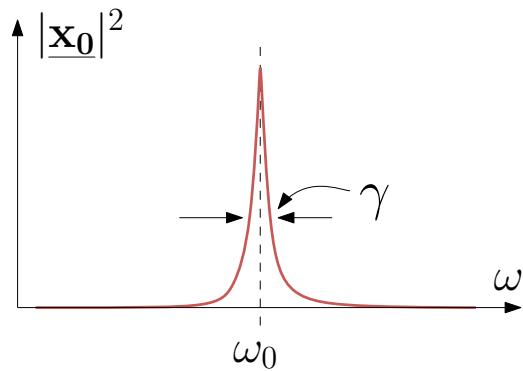
The squared amplitude of the electron oscillation then reads

$$|\underline{\mathbf{x}}_0|^2 = \frac{e^2}{m_e^2} \frac{|\underline{\mathbf{E}}_0|^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2 \omega^2}.$$

and exhibits a resonance at $\omega_{\text{res}} = \sqrt{\omega_0^2 - \gamma^2/2} \simeq \omega_0$ for which the amplitude

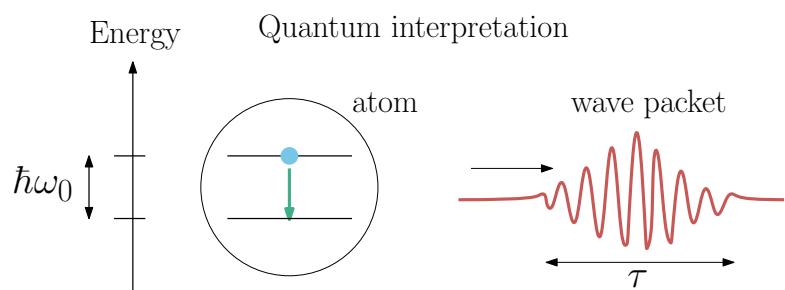
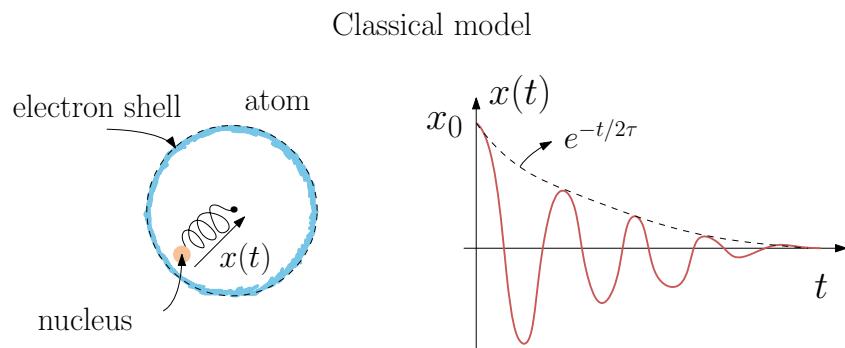
$$|\underline{\mathbf{x}}_0(\omega_{\text{res}})| = \frac{e}{m_e} \frac{|\underline{\mathbf{E}}_0|}{\gamma \sqrt{\omega_0^2 - \gamma^2/4}} \simeq \frac{e}{m_e} \frac{|\underline{\mathbf{E}}_0|}{\gamma \omega_0}$$

is very large (this would correspond to a transition in quantum mechanics). The width of the resonance is given by γ which, for an isolated atom ($\gamma = 10^8 \text{ s}^{-1}$) and a visible electromagnetic wave ($\omega_0 \sim 3 \times 10^{15} \text{ s}^{-1}$) is such that $\gamma/\omega_0 \sim 10^{-8}$. In that case the amplitude as a function of frequency has approximately a Cauchy-Lorentzian shape as shown in the figure below



Remarks

- The damping frequency γ is related to the typical time $\tau = 1/\gamma$ that it takes for the amplitude of an oscillator to decay by a factor of e , and it is related to the energy lost in emitted radiation by the oscillating dipole. In quantum mechanics, τ has a different interpretation: an electron in an excited state will spontaneously decay to the lowest energy state within a time τ , emitting a wave packet of frequency ω_0 and of typical duration τ . The frequency ω_0 of the emitted wave is related to the energy difference ΔE between the excited and lowest energy state, by $\Delta E = \hbar\omega_0$, where \hbar is the reduced Planck constant.



- This simple model, although representing an atom, can be extended to describe more complex systems such as molecules (whose vibrational modes are quantized as well and therefore exhibit resonant frequencies) or crystals. As such, gases, liquids or solid media can be modeled by an assembly of harmonic oscillators. Typically, the linewidth γ of the resonances in matter is dominated by collisions between the atoms and not by spontaneous emission, as in the case of isolated atoms.

11.3.2 The complex permittivity

If n is the density of atomic (or molecular) dipoles in the medium, each one having Z electrons and a single resonance ω_0 , the polarization vector reads

$$\underline{\mathbf{P}} = n(-Z e \underline{\mathbf{x}}) = \frac{-n Z e^2}{m_e} \frac{\underline{\mathbf{E}}}{\omega^2 - \omega_0^2 + i\gamma\omega} = (\underline{\epsilon}(\omega) - \epsilon_0) \underline{\mathbf{E}},$$

so that the complex permittivity is given by

$$\underline{\epsilon}(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \right), \quad (11.14)$$

where $\omega_p = \sqrt{Zne^2/m_e\epsilon_0}$ is the plasma frequency. Separating explicitly its real and imaginary parts gives

$$\underline{\epsilon}(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} + i \frac{\omega_p^2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \right). \quad (11.15)$$

Note that in the limit of very high frequencies, the medium becomes transparent, $\lim_{\omega \rightarrow \infty} \epsilon(\omega) = \epsilon_0$, whereas for very low frequencies we get an expression for the static permittivity,

$$\lim_{\omega \rightarrow 0} \epsilon(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2}{\omega_0^2} \right) = \epsilon.$$

11.3.3 The complex refractive index

Finally, the complex index of refraction \underline{n} , which fully characterizes the propagation of a wave in a medium, is complex whenever $\underline{\epsilon}(\omega)$ has a non-negligible imaginary part (i.e. close to resonant frequencies). In this case, it writes

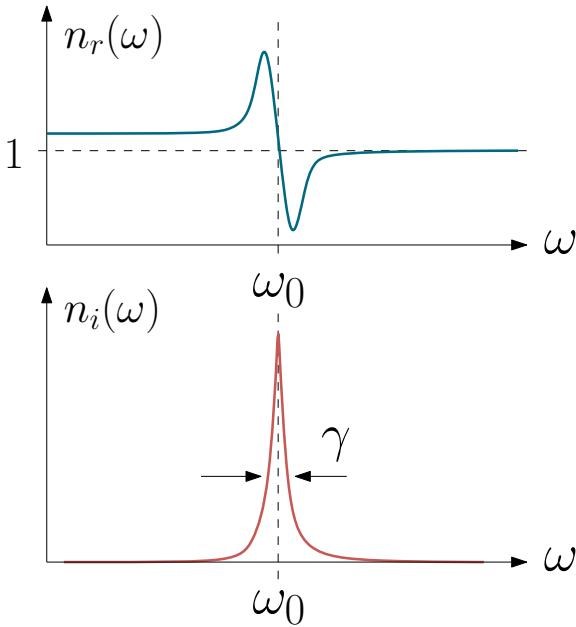
$$\underline{n}(\omega) = \sqrt{\frac{\underline{\epsilon}(\omega)}{\epsilon_0}} = n_R(\omega) + i n_I(\omega).$$

For a diluted medium, one has $\omega_p \ll \omega_0$. We can then use the expansion

$$\sqrt{\frac{\underline{\epsilon}(\omega)}{\epsilon_0}} = \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \right)^{\frac{1}{2}} \approx 1 + \frac{\omega_p^2}{2(\omega_0^2 - \omega^2 - i\gamma\omega)}$$

and so,

$$\underline{n}(\omega) \approx \underbrace{1 + \frac{1}{2} \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}_{n_R} + i \underbrace{\frac{1}{2} \frac{\omega_p^2\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}_{n_I} \quad (11.16)$$



Remarks

- Away from any resonance, $n_I \ll n_R$ so that absorption may be neglected and the refractive index may be considered as a pure real number, $n \approx n_R$.
- The very large change of index of refraction n_R close to a resonance will cause a large dispersion (see section 11.10), i.e., different wavelengths travel at very different speeds near the resonance.
- Note that for frequencies slightly above a very strong resonance, the real part of the refractive index could be less than one. This does not mean, however, than an electromagnetic wave centered around this frequency will travel faster than c . As will be shown in section 11.10, any physical wave is represented by a wave packet, that is, a superposition of plane waves. The speed at which such a wave packet moves is not given by $\omega/k(\omega)$ but instead by $\partial\omega/\partial k$, also called group velocity.
- Finally, in real materials there are several types of resonances distributed in the whole electromagnetic spectrum (vibrations fall in the infrared, electronic resonances in atoms typically in the UV or visible range, etc). A realistic model including more than one resonant frequency reads

$$\underline{n}^2(\omega) = 1 + \omega_p^2 \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - i\gamma_j\omega},$$

where f_j , called the oscillator strength, represents the relative weight of the frequency ω_j which can be properly determined within the framework of quantum mechanics.

11.4 The laws of reflection and refraction

11.4.1 Introduction

The phenomenon of reflection and refraction of light has been known since ancient times. When light hits the interface between two media, it splits into a reflected and a refracted beam, the latter being *bent* with respect to the initial propagation direction. When one of the two interfaces is a metal, most of the light is reflected and if the surface is smooth enough, it acts like a mirror. Refraction, on the other hand, is observed in transparent materials such as glass. Tailoring of the interface profile is used to create lenses and optical systems, a subject that was treated during the course of electromagnetism and light (PHY 104).

Refraction being generally a wavelength-dependent phenomenon, it is also at the origin of rainbows and of the dispersion of sunlight by a prism. Exploiting the laws of reflection and refraction has allowed us to conceive clever systems of high relevance in modern optics such as dielectric mirrors with ultra high reflectivity, interferential filters, or resonant optical cavities for spectroscopy or the fabrication of lasers. In this section, we will demonstrate that the laws of reflection and refraction are a consequence of the continuity conditions of the electromagnetic fields derived from Maxwell's equations.

11.4.2 Continuity equations at an interface

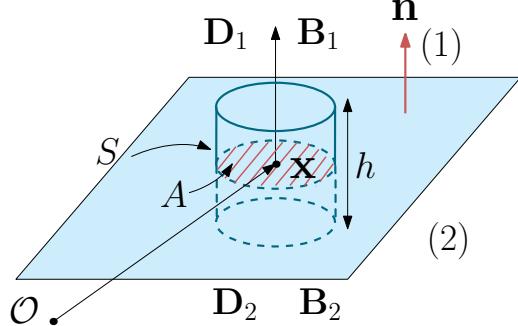
In this section we will derive from Maxwell's equations the continuity conditions for the electromagnetic fields at an interface separating two media. Of course, no abrupt interface exists at the atomic scale, but it is a convenient way of dealing with the averaged electromagnetic fields in matter. Here, we suppose that the fields are averaged over a typical length scale l such that $a \ll l \ll \lambda$, where $a \sim 0.1 \text{ nm}$ is the typical size of an atom and λ is the wavelength of the electromagnetic wave. When this is possible, matter can be modeled as a continuous medium characterized by a refractive index n . Note however that when $\lambda \sim a$, there is no intermediate length scale l and this averaging process is no longer possible. In this regime, one cannot describe the interaction of waves with matter in terms of reflection and refraction. This is the case, for example, of X-rays which are known to be diffracted when incident on crystals.

- **NORMAL COMPONENTS OF \mathbf{B} AND \mathbf{D} AT AN INTERFACE**

In their integral form, the Gauss-Maxwell and Thomson-Maxwell equations for the averaged fields read

$$\oint_S \mathbf{D} \cdot \mathbf{n} dS = Q(S) \quad \text{and} \quad \oint_S \mathbf{B} \cdot \mathbf{n} dS = 0.$$

Consider now an interface separating two media 1 and 2, and let \mathbf{n} be the normal to the interface at point \mathbf{x} , directed from 2 to 1. We consider an infinitesimal, closed cylindrical surface of height h and transverse area A , as shown in the figure below.



When $h \rightarrow 0$, the flux integral over S reduces to the flux through the caps just above and just below the interface

$$\iint_S \mathbf{D} \cdot \mathbf{n} dS = A(\mathbf{D}_1(\mathbf{x}) - \mathbf{D}_2(\mathbf{x})) \cdot \mathbf{n},$$

$$\iint_S \mathbf{B} \cdot \mathbf{n} dS = A(\mathbf{B}_1(\mathbf{x}) - \mathbf{B}_2(\mathbf{x})) \cdot \mathbf{n}.$$

The zero-divergence of the magnetic field ensures the continuity of its component normal to the interface,

$$\mathbf{n} \cdot (\mathbf{B}_1 - \mathbf{B}_2) |_{\mathbf{x}} = 0, \quad (11.17)$$

whereas for the displacement vector \mathbf{D} the flux through S may differ from zero if the interface is charged with a surface density σ , in which case the charge $Q(S)$ enclosed by the cylinder is $A\sigma$ and we conclude that the normal component of \mathbf{D} presents a discontinuity whenever $\sigma \neq 0$:

$$\mathbf{n} \cdot (\mathbf{D}_1 - \mathbf{D}_2) |_{\mathbf{x}} = \sigma(\mathbf{x}). \quad (11.18)$$

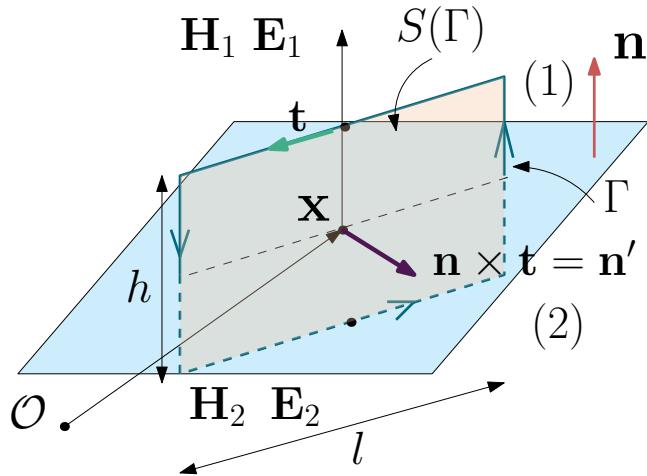
• TANGENTIAL COMPONENTS OF \mathbf{E} AND \mathbf{H} AT AN INTERFACE

Now we recall the Faraday-Maxwell and the Ampère-Maxwell laws for the averaged fields in their integral form

$$\oint_{\Gamma} \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \iint_{S(\Gamma)} \mathbf{B} \cdot \mathbf{n} dS,$$

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{l} = \iint_{S(\Gamma)} \mathbf{J} \cdot \mathbf{n} dS + \frac{d}{dt} \iint_{S(\Gamma)} \mathbf{D} \cdot \mathbf{n} dS.$$

We can choose for Γ an infinitesimal, closed rectangular path centered at \mathbf{x} , of height h and length l and let $S(\Gamma)$ be the planar surface enclosed by Γ , whose normal \mathbf{n}' is tangent to the interface, as shown in the figure below.



In the limit $h \rightarrow 0$, the circulation along Γ reduces to the integrals along the horizontal segments parallel to the surface, so that

$$\oint_{\Gamma} \mathbf{E} \cdot d\mathbf{l} = l \mathbf{t} \cdot (\mathbf{E}_1(\mathbf{x}) - \mathbf{E}_2(\mathbf{x})),$$

$$\oint_{\Gamma} \mathbf{H} \cdot d\mathbf{l} = l \mathbf{t} \cdot (\mathbf{H}_1(\mathbf{x}) - \mathbf{H}_2(\mathbf{x})).$$

where \mathbf{t} is a unit vector tangent to the interface along the horizontal path just above it. The flux integrals of \mathbf{B} and \mathbf{D} tend to zero as $h \rightarrow 0$ (hence, the surface area tends to zero). In consequence, the tangential component \mathbf{E}^{\parallel} of the electric field is continuous at any point \mathbf{x} of the interface :

$$\boxed{\mathbf{E}_2^{\parallel}(\mathbf{x}) - \mathbf{E}_1^{\parallel}(\mathbf{x}) = 0.} \quad (11.19)$$

For the magnetic excitation \mathbf{H} , however, a non-zero flux of the current density over $S(\Gamma)$ is possible if a current is confined at the interface, in which case $\iint_{S(\Gamma)} \mathbf{J} \cdot \mathbf{n} dS = l \mathbf{J}_s \cdot \mathbf{n}'$ where \mathbf{J}_s is the current density (in A m^{-1}) at the interface. Since $\mathbf{t} = \mathbf{n}' \times \mathbf{n}$ and using $(\mathbf{n}' \times \mathbf{n}) \cdot (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{n}' \cdot (\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2))$ we conclude that

$$\boxed{\mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) |_{\mathbf{x}} = \mathbf{J}_s.} \quad (11.20)$$

11.4.3 Reflection and refraction of an electromagnetic wave at an interface

When an incident beam of light arrives on a flat interface between two isotropic media, it gives birth to a reflected and a transmitted beam. Here, we will show that the existence of these waves results from the continuity conditions of the electromagnetic fields at the interface. When the interface is not flat, one can always apply the laws of reflection and refraction locally, at every point of the interface. Recall

that a sinusoidal electromagnetic plane wave in matter can be written in its complex representation as

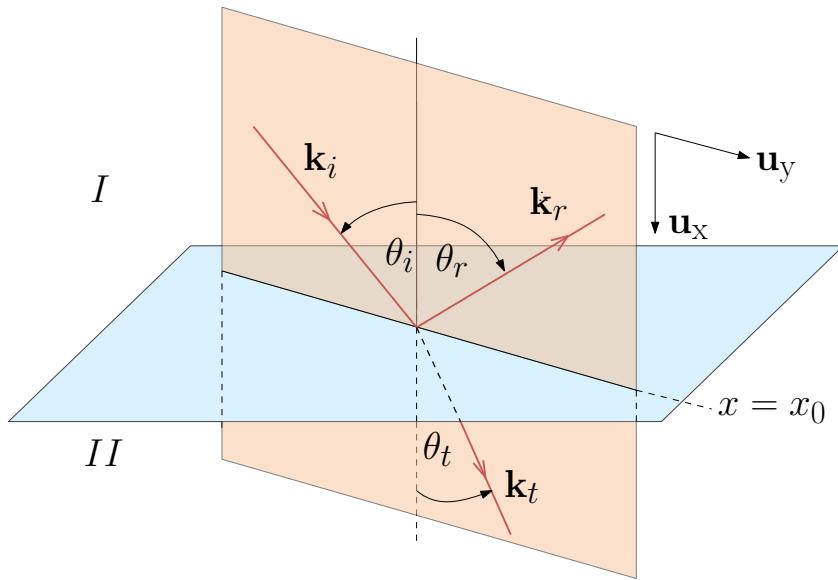
$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

where the wavevector writes $\mathbf{k} = k\mathbf{n}$, with \mathbf{n} a unit vector representing the direction of propagation of the wave and k the wavenumber. We limit the discussion to the case of dielectric media without absorption. Hence, the wavenumber k is real-valued and is related to the refractive index $n(\omega)$ by the dispersion relation

$$k(\omega) = \frac{\omega}{c} n(\omega),$$

where $n(\omega)$ is also real valued and may depend on the wave frequency ω . Let us now consider the situation in which a sinusoidal plane wave encounters an interface between two media, represented by the half-spaces $I : \{x < x_0\}$ and $II : \{x > x_0\}$. Let Oxy be the plane containing both the normal to the interface and the direction of propagation of the incident wave. At $x = x_0^-$ (region I), the total electric field is the superposition of an incident wave ($\underline{\mathbf{E}}_i$) and a wave reflected off the interface ($\underline{\mathbf{E}}_r$). The incident wave, whose direction of propagation is given by \mathbf{k}_i and forms an angle θ_i with respect to the normal to the interface, is given at $x = x_0^-$ by

$$\underline{\mathbf{E}}_i(x_0^-, y, t) = \underline{\mathbf{E}}_{i0} e^{i(k_i \cos \theta_i x_0 + k_i \sin \theta_i y - \omega_i t + \phi_i)}$$



and the reflected wave writes

$$\underline{\mathbf{E}}_r(x_0^-, y, t) = \underline{\mathbf{E}}_{r0} e^{i(-k_r \cos \theta_r x_0 + k_r \sin \theta_r y - \omega_r t + \phi_r)},$$

where the wave vector \mathbf{k}_r indicates the propagation direction forming an angle θ_r with respect to the normal to the interface. Finally, in medium II at $x = x_0^+$ there will be a transmitted wave

$$\underline{\mathbf{E}}_t(x_0^+, y, t) = \underline{\mathbf{E}}_{t0} e^{i(k_t \cos \theta_t x_0 + k_t \sin \theta_t y - \omega_t t + \phi_t)},$$

where k_t and θ_t denote the wavenumber and the angle formed by the transmitted wave propagation direction with respect to the normal to the interface. The total electric field $\underline{\mathbf{E}}_i + \underline{\mathbf{E}}_r$ in region I and

the field $\underline{\mathbf{E}}_t$ in region II are related by the continuity conditions (11.18) and (11.19). In particular, the tangential component of the electric field is continuous at $x = x_0$ for arbitrary y and t . This is only possible if the argument of the complex exponentials are the same for all the three waves³. At $t = 0$ and $y = 0$, this writes

$$k_i \cos \theta_i x_0 + \phi_i = -k_r \cos \theta_r x_0 + \phi_r = k_t \cos \theta_t x_0 + \phi_t \quad (11.21)$$

At $t = 0$ and y arbitrary, and using Eq. (11.21)

$$k_i \sin \theta_i y = k_r \sin \theta_r y = k_t \sin \theta_t y. \quad (11.22)$$

Finally, at $y = 0$, t arbitrary and using Eqs. (11.21) and (11.22)

$$\omega_i t = \omega_r t = \omega_t t \quad (11.23)$$

From these conditions we conclude that :

1. The frequency ω of the reflected and transmitted waves is the same as that of the incident one (see Eq.(11.23))

$$\omega = \omega_i = \omega_r = \omega_t.$$

The incident wave oscillating at ω forces the atomic dipoles to oscillate and to re-emit light at the same frequency.

2. Since the reflected wave satisfies the same dispersion relation as the incident one, we have $k_i = k_r = n(\omega)\omega/c$ and we conclude from Eq. (11.22) that the reflected wave forms an angle

$$\boxed{\theta_r = \theta_i} \quad (11.24)$$

with respect to the normal, and so the reflected wavevector is the symmetric of the incident wavevector with respect to the interface. Equation (11.21) then rewrites

$$k_i \cos \theta_i x_0 + \phi = -k_i \cos \theta_i x_0 + \phi_r = k_t \cos \theta_t x_0 + \phi_t,$$

which yields ϕ_r and ϕ_t as functions of k_i , θ_i and ϕ_i .

3. From Eq. (11.22), we obtain the Snell-Descartes law

$$\boxed{k_i \sin \theta_i = k_t \sin \theta_t} \quad (11.25)$$

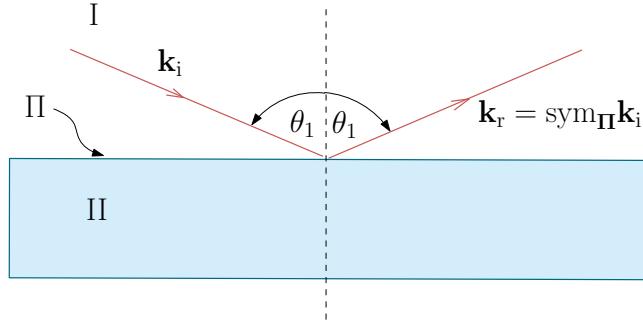
which expresses the fact that the component of the wavevector that is tangent to the interface is the same for the incident, reflected and transmitted waves. Snell-Descartes's law can be rewritten in terms of the refractive index as discussed below.

³Note that we have already assumed that the reflected and transmitted wavevectors lie both in the plane of incidence. If this were not the case, it would be impossible to fulfill the continuity of the tangential component of \mathbf{E} for every point at the interface.

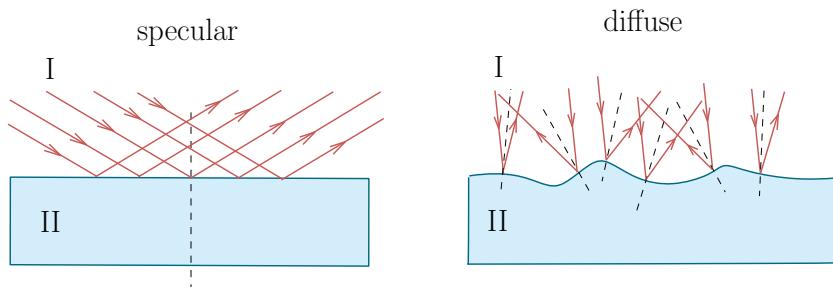
- REFLECTION AND REFRACTION LAWS

Imposing the continuity of the tangential component of the electric field at an interface requires the existence of a reflected and a transmitted waves whose wavevectors lie in the same plane and are given by the following reflection and refraction laws:

- **Reflection law:** the reflected wavevector is the symmetric of the incident wavevector with respect to the interface II. They both form the same angle θ_1 with respect to the normal to the interface.

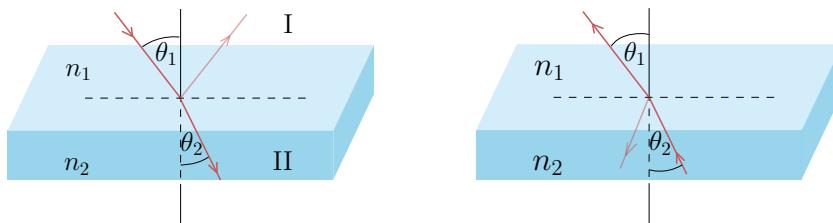


Specular reflection is the term used to distinguish the reflection from a very flat surface (such as a mirror) from the diffuse reflection observed on irregular surfaces (most of everyday objects). Specular reflection allows us to form images of objects since all the points on the surface reflect light in the same direction, whereas for diffuse reflection light is scattered in all directions since the normal to the interface changes rapidly from one point to the other. It is the latter kind of reflection that allows us to observe the world around us.



- **Snell-Descartes's law:** If a monochromatic plane wave coming from medium I is incident on the interface separating media I and II, of refractive index n_1 and n_2 , respectively, then the projection along the normal of the transmitted wavevector in medium II is characterized by an angle θ_2 given by

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (11.26)$$



One consequence of the law of refraction is the principle of backward propagation of light: if light is incident from medium II at an angle θ_2 , then the refracted wave in medium I forms an angle θ_1 with respect to the normal. More generally, for non absorbing dielectric media, every path followed by a beam of light in one direction can be traveled in the opposite sense.

• THE CASE OF TOTAL REFLECTION

In dielectrics far from any absorption resonance, the refractive index is real, and the Snell-Descartes law becomes

$$\sin \theta_2 = \frac{n_1}{n_2} \sin \theta_1.$$

A particular case arises whenever medium 1 has a refractive index larger than that of medium 2, i.e., when $n_1 > n_2$. In that case, a real solution for θ_2 exists only if

$$\frac{n_1}{n_2} \sin \theta_1 \leq 1,$$

which is satisfied when the incident angle does not exceed a critical value θ_1^* given by

$$\theta_1^* = \sin^{-1} \left(\frac{n_2}{n_1} \right).$$

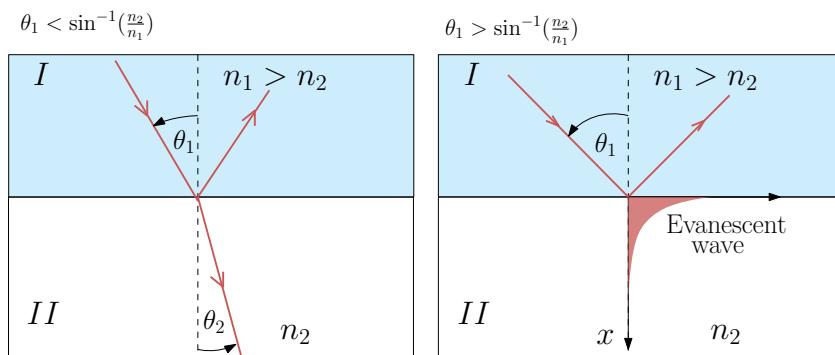
Then we have a reflected and a transmitted beam. However, there is no real solution for θ_2 if $\theta_1 > \theta_1^*$. In this case, θ_2 is a complex number such that

$$\cos \theta_2 = i \sqrt{\underbrace{\frac{n_1}{n_2} \sin^2 \theta_1 - 1}_{>0}}$$

so that the transmitted wave adopts the form of an evanescent wave propagating along the surface (Oy direction) and decaying exponentially as a function of x inside region II:

$$\underline{\mathbf{E}}_t(x, y, t) = \underline{\mathbf{E}}_{t0} e^{i(\frac{\omega}{c} n_1 \sin \theta_1 y - \omega t + \phi_t)} e^{-\frac{\omega}{c} \sqrt{\frac{n_1}{n_2} \sin^2 \theta_1 - 1} x}$$

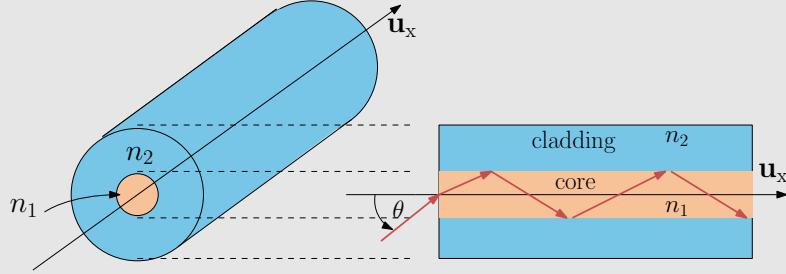
This case corresponds to a total reflection of the incident beam.



11.5 Dielectric waveguides

Example 11.2 - The optical fiber

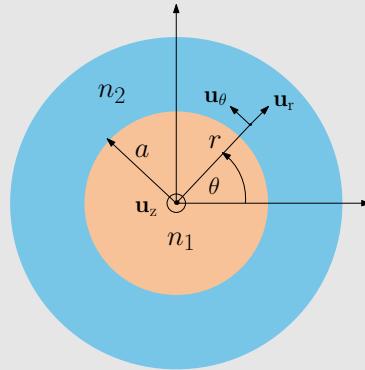
An optical fiber is made of a cylindrical dielectric medium of refractive index n_1 , called the core, in which light propagates. The core is surrounded by a second, concentric cylinder of refractive index n_2 . The indexes are chosen such that $n_2 < n_1$ so that light can be confined in the core by assuring total reflection at the core-cladding interface.



- a. Suppose that an incident plane wave arrives from air ($n = 1$) and enters the core with an angle θ with respect to the axis of the fiber. What is the maximum value of θ that will assure total reflection at the core-cladding interface?
- b. We will treat the problem differently by looking for a solution to the propagation equation for which the field is confined inside the fiber. We define region 1 as $r < a$ (core) and region 2 as $a < r$ (cladding, supposed to have an infinite extension). Suppose a monochromatic wave propagating along the z -axis in region i , which in cylindrical coordinates writes

$$\underline{\mathbf{E}}_i(r, \theta, z, t) = E_i(r, \theta) e^{i(kz - \omega t)} \mathbf{u}$$

where $\mathbf{u} \perp \mathbf{u}_z$ is the wave polarization.



What is the equation that $E_i(r, \theta)$ must satisfy if medium i has a constant refractive index n_i ?

- c. We write down the solution under the form $E_i(r, \theta) = \mathcal{E}_i(r) e^{\pm i l \theta}$ where $l \in \mathbb{N}$. Show that \mathcal{E}_i is therefore a solution of

$$r^2 \frac{d^2 \mathcal{E}_i}{dr^2} + r \frac{d\mathcal{E}_i}{dr} + \left(\frac{r^2}{a^2} q_i(r)^2 - l^2 \right) \mathcal{E}_i = 0$$

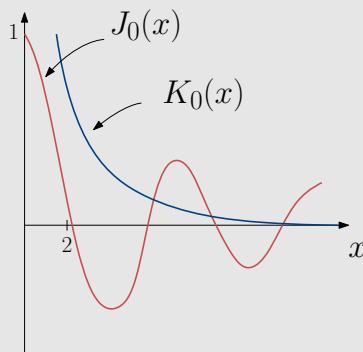
and determine $q_i(r)$.

d. The solutions of the equation

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + \left(q^2 \frac{x^2}{a^2} - l^2 \right) y = 0$$

are called Bessel functions. Here, we will only consider Bessel functions that do not diverge at $x = 0$ or when $x \rightarrow \infty$, which are of two kinds:

- if $q^2 > 0$, the solution writes in terms of the Bessel function of the first kind, $y(x) = J_l(qx/a)$. It oscillates while decaying slowly to zero as x increases.
- if $q^2 < 0$, the solution writes in terms of the Bessel function of the second kind, $y(x) = K_l(hx/a)$ where $h^2 = -q^2 > 0$, which decays exponentially as x increases.

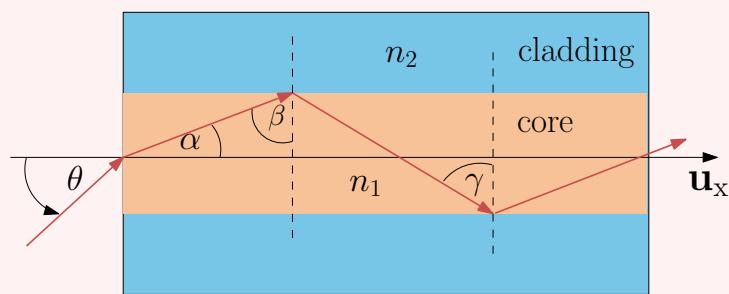


On this basis, what are the possible values of k for which the wave is confined in the core of the fiber? Show that this is equivalent to the condition found in the first question with the Snell-Descartes law for the maximum angle of entrance in the fiber.

Solution

- a.** If the wave enters the core forming an angle θ with respect to the axis of the fiber (and therefore with respect to the air-core interface), then the transmitted wave forms an angle α with the axis given by the Snell-Descartes law (11.26):

$$\sin \theta = n_1 \sin \alpha.$$



The wave will encounter the core-cladding interface forming an angle $\beta = \pi/2 - \alpha$ with respect to the normal. The reflection law then assures that $\gamma = \beta$ so that all the subsequent reflections at the core-cladding interface occur at the same incident angle. Light will be confined inside the core if all these reflections are total, that is, if β is larger than the critical angle $\beta^* = \sin^{-1} \frac{n_2}{n_1}$, and so

$$\beta = \frac{\pi}{2} - \alpha \geq \sin^{-1} \frac{n_2}{n_1}$$

and in terms of θ

$$\frac{\pi}{2} - \sin^{-1} \left(\frac{\sin \theta}{n_1} \right) \geq \sin^{-1} \frac{n_2}{n_1}.$$

Equivalently,

$$\sin \left(\frac{\pi}{2} - \sin^{-1} \left(\frac{\sin \theta}{n_1} \right) \right) = \underbrace{\cos \left(\sin^{-1} \left(\frac{\sin \theta}{n_1} \right) \right)}_{\sqrt{1 - \sin^2(\sin^{-1}(\sin \theta / n_1))}} \geq \frac{n_2}{n_1}.$$

Finally,

$$\sqrt{1 - \frac{\sin^2 \theta}{n_1^2}} \geq \frac{n_2}{n_1} \rightarrow \sin \theta \leq \sqrt{n_1^2 - n_2^2},$$

leading to

$$\theta \leq \sin^{-1} \left(\sqrt{n_1^2 - n_2^2} \right).$$

- b.** In region i (with $i = \text{core, cladding}$), $\underline{\mathbf{E}}_i$ must be a solution of d'Alembert's wave equation for a monochromatic wave

$$\nabla^2 \underline{\mathbf{E}}_i + \frac{n_i^2 \omega^2}{c^2} \underline{\mathbf{E}}_i = 0,$$

where the second order time derivative was directly replaced by $-\omega^2$ using complex notations.

Note that this is true provided the refractive index in each medium is constant in space. Indeed when the refractive index is inhomogeneous, there is a non-zero polarization charge density. The right-hand side term in the wave equation is non-zero, making possible the existence of longitudinal waves, for which the electric field may have a non-zero component along \mathbf{u}_z .

Replacing the Laplacian in cylindrical coordinates and injecting a wave of the form $\underline{\mathbf{E}}_i(r, \theta, z, t) = E_i(r, \theta) e^{i(kz - \omega t)} \mathbf{u}$, where \mathbf{u} is a constant transverse unit vector that represents the polarization direction of the wave, we obtain

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial E_i(r, \theta)}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 E_i(r, \theta)}{\partial \theta^2} = k^2 E_i(r, \theta) - \frac{n_i^2(r) \omega^2}{c^2} E_i(r, \theta).$$

c. Inserting a solution of the form $\mathcal{E}_i(r)e^{\pm il\theta}$ into the latter equation yields

$$\frac{1}{r} \left(\frac{d\mathcal{E}_i}{dr} + r \frac{d^2\mathcal{E}_i}{dr^2} \right) e^{\pm il\theta} - \left(\frac{l^2}{r^2} + k^2 \right) \mathcal{E}_i(r) e^{\pm il\theta} = -\frac{n_i^2(r)\omega^2}{c^2} \mathcal{E}_i(r) e^{\pm il\theta}$$

and we find:

$$r \frac{d\mathcal{E}_i}{dr} + r^2 \frac{d^2\mathcal{E}_i}{dr^2} - \left(l^2 + r^2 k^2 \right) \mathcal{E}_i(r) = -\frac{r^2 n_i^2(r)\omega^2}{c^2} \mathcal{E}_i(r)$$

which can be rewritten as

$$r \frac{d\mathcal{E}_i}{dr} + r^2 \frac{d^2\mathcal{E}_i}{dr^2} + \left(\frac{r^2}{a^2} q_i^2(r) - l^2 \right) \mathcal{E}_i(r) = 0$$

with

$$q_i^2(r) = a^2 \left(\frac{n_i^2 \omega^2}{c^2} - k^2 \right).$$

d. A wave confined in the core of the fiber is obtained if $q_1^2 > 0$ (the field inside the core behaves like a Bessel function of the first kind) and if $q_2^2 < 0$ (the field in the cladding behaves like a Bessel function of the second kind, and therefore decays exponentially as r increases). This means

$$\frac{n_1^2 \omega^2}{c^2} - k^2 > 0 \text{ and } \frac{n_2^2 \omega^2}{c^2} - k^2 < 0,$$

and so the wavenumber along Oz must satisfy

$$\frac{n_2^2 \omega^2}{c^2} < k^2 < \frac{n_1^2 \omega^2}{c^2}.$$

Recalling that $k_0 = \omega/c$ is the wavenumber associated with the wave in vacuum, finally the condition for a guided mode writes

$$n_2^2 k_0^2 < k^2 < n_1^2 k_0^2.$$

This guided electric field is the superposition of plane waves propagating in the core and reflected at the core-cladding interface. As such, the wavenumber k along z corresponds to $k = n_1 k_0 \cos \alpha = n_1 k_0 \sqrt{1 - \sin^2 \alpha}$ where α is the angle formed by the plane waves with respect to the fiber axis. The Snell-Descartes law gives $\sin \alpha = \frac{\sin \theta}{n_1}$ and so we find

$$n_2^2 k_0^2 < k_0^2 (n_1^2 - \sin^2 \theta) < n_1^2 k_0^2,$$

leading to

$$0 < \sin^2 \theta < n_1^2 - n_2^2.$$

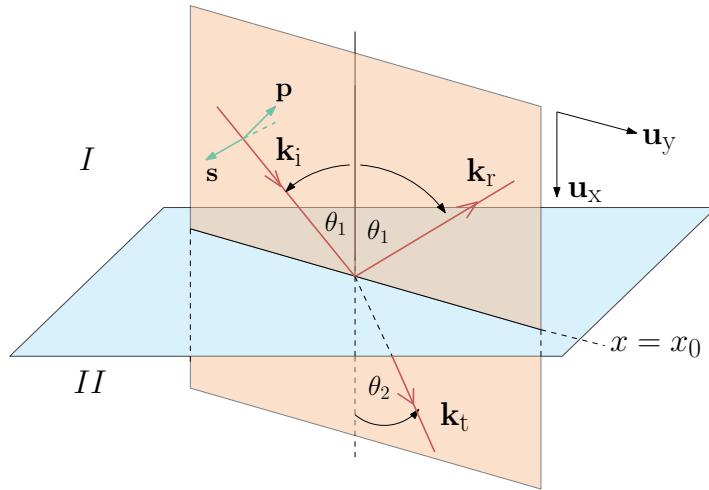
11.6 Fresnel's reflection and transmission coefficients

The reflection and refraction laws, coming from the continuity of the tangential component of the electric field at an interface, give us the propagation direction of the reflected and transmitted waves at an interface. They do not tell us, however, what fraction of the incident intensity is reflected and what fraction is transmitted. We know for example that a metallic mirror will reflect most of the incident light, whereas a transparent dielectric such as a window made of glass will reflect only a small percentage of the incident intensity; most of it will be transmitted through it. In order to predict the distribution of the energy flux between reflected and transmitted waves, we need to impose the other continuity conditions for the electromagnetic fields.

For a given polarization of the electric field, let us then write the incident, reflected and transmitted waves at an interface $x = x_0$ as follows

$$\begin{aligned} \text{incident wave: } & \underline{\mathbf{E}}_i(x_0^-, y, t) = \underline{\mathbf{E}}_{i0} e^{i(\mathbf{k}_i \cdot \mathbf{x} - \omega t + \phi_i)} \\ \text{reflected wave: } & \underline{\mathbf{E}}_r(x_0^-, y, t) = \underline{\mathbf{E}}_{r0} e^{i(\mathbf{k}_r \cdot \mathbf{x} - \omega t + \phi_r)} \\ \text{transmitted wave: } & \underline{\mathbf{E}}_t(x_0^+, y, t) = \underline{\mathbf{E}}_{t0} e^{i(\mathbf{k}_t \cdot \mathbf{x} - \omega t + \phi_t)} \end{aligned}$$

where we know from the continuity of the tangential component of the electric field at the boundary $x = x_0$ that the arguments of the complex exponentials are the same for every y and t . This implies in particular the laws of reflection and refraction for the angles that the wavevectors form with the normal to the interface.



If we now impose the continuity conditions (11.17), (11.18) and (11.20) at the interface, we may compute the amplitude of the reflected and transmitted waves in terms of the incident one, which will allow us to define a reflection, r , and a transmission coefficient t , also called Fresnel's coefficients, as follows

$$\underline{\mathbf{E}}_{r0} = r \underline{\mathbf{E}}_{i0}, \quad (11.27)$$

$$\underline{\mathbf{E}}_{t0} = t \underline{\mathbf{E}}_{i0}. \quad (11.28)$$

Remarkably, these coefficients depend on the polarization of the incident wave, which in the most general case can be decomposed as a linear superposition of an *s* polarized wave⁴ (electric field perpendicular to the plane of incidence) and a *p* polarized wave (electric field parallel to, or in the incidence plane). The general reflection and transmission coefficients will be therefore a linear combination of the ones for *s*- and *p*-polarization that we separately calculate below.

11.6.1 Fresnel's coefficients for perpendicular *s*-polarization

An *s*-polarized wave is such that its polarization is perpendicular to the plane of incidence, and so its electric field is tangential to the interface. In the figure below, this corresponds to an electric field along Oz , so that at the interface, the amplitude of the electric field writes

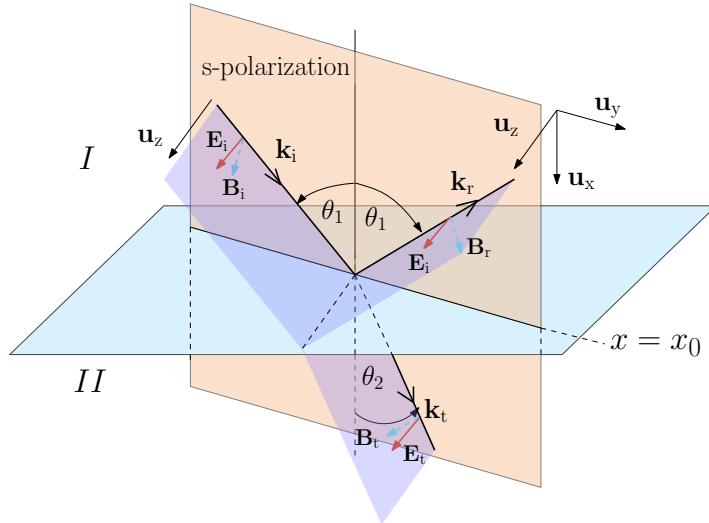
$$\begin{aligned} \text{incident wave: } & \underline{\mathbf{E}}_{i0} = E_{i0} \mathbf{u}_z \\ \text{reflected wave: } & \underline{\mathbf{E}}_{r0} = E_{r0} \mathbf{u}_z \\ \text{transmitted wave: } & \underline{\mathbf{E}}_{t0} = E_{t0} \mathbf{u}_z \end{aligned}$$

and since according to Eq. (11.19), the tangential component of the electric field is continuous, we have

$$E_{i0} + E_{r0} = E_{t0} \quad (11.29)$$

or equivalently

$$1 + r_s = t_s. \quad (11.30)$$



On the other hand, the incident, reflected and transmitted magnetic fields can be obtained from $\underline{\mathbf{B}} = \frac{\mathbf{k}}{\omega} \times \underline{\mathbf{E}}$, which yields:

$$\begin{aligned} \underline{\mathbf{B}}_{i0} &= \underbrace{\frac{n_1}{c}}_{k/\omega} (\cos \theta_1 \mathbf{u}_x + \sin \theta_1 \mathbf{u}_y) \times \underline{\mathbf{E}}_{i0} \\ &= \frac{n_1}{c} \underline{\mathbf{E}}_{i0} (\sin \theta_1 \mathbf{u}_x - \cos \theta_1 \mathbf{u}_y), \end{aligned}$$

⁴from the German word *senkrecht*, meaning *perpendicular*.

and similarly

$$\begin{aligned}\underline{\mathbf{B}}_{r0} &= \frac{n_1}{c} \underline{\mathbf{E}}_{r0} (\sin \theta_1 \mathbf{u}_x + \cos \theta_1 \mathbf{u}_y), \\ \underline{\mathbf{B}}_{t0} &= \frac{n_2}{c} \underline{\mathbf{E}}_{t0} (\sin \theta_2 \mathbf{u}_x - \cos \theta_2 \mathbf{u}_y).\end{aligned}$$

It is easy to check that imposing the continuity condition (11.17) for the normal component of the magnetic field gives the same condition as Eq. (11.30). On the other hand, Eq. (11.20) imposes the continuity of the tangential component (along Oy) of \mathbf{H} , since the media are non-magnetic and no current flows at the interface:

$$-\underline{E}_0 n_1 \cos \theta_1 + \underline{E}_{r0} n_1 \cos \theta_1 = -\underline{E}_{t0} n_2 \cos \theta_2. \quad (11.31)$$

or equivalently

$$n_1 \cos \theta_1 (1 - r_s) = t_s n_2 \cos \theta_2. \quad (11.32)$$

Combining Eqs. (11.30) and (11.32) allows us to find Fresnel's coefficients giving the reflected and transmitted amplitudes in terms of the incident one for an s -polarized wave:

$$\begin{aligned}r_s &= \frac{\underline{\mathbf{E}}_{r0}}{\underline{\mathbf{E}}_0} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2}, \\ t_s &= \frac{\underline{\mathbf{E}}_0^t}{\underline{\mathbf{E}}_0} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + n_2 \cos \theta_2}.\end{aligned} \quad (11.33)$$

11.6.2 Fresnel's coefficients for parallel p -polarization

A p -polarized wave is such that its polarization is in the plane of incidence, and so its magnetic field is tangential to the interface. The amplitude of the electric field writes

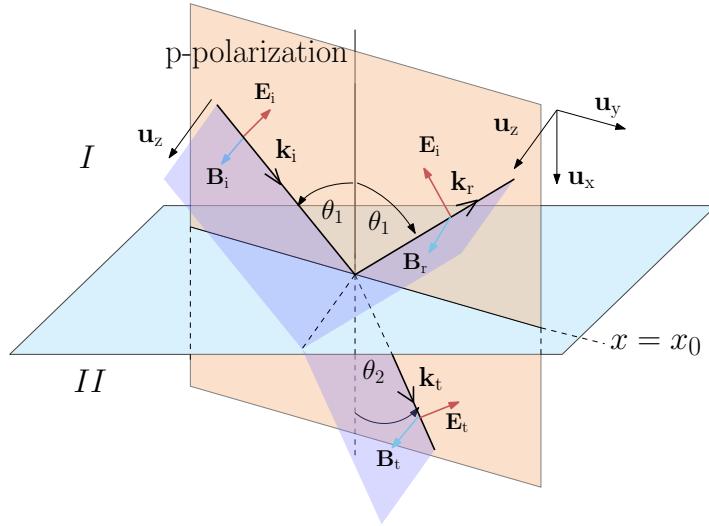
$$\begin{aligned}\text{incident wave: } \underline{\mathbf{E}}_{i0} &= \underline{\mathbf{E}}_0 (-\sin \theta_1 \mathbf{u}_x + \cos \theta_1 \mathbf{u}_y) \\ \text{reflected wave: } \underline{\mathbf{E}}_{r0} &= \underline{\mathbf{E}}_0 (-\sin \theta_1 \mathbf{u}_x - \cos \theta_1 \mathbf{u}_y) \\ \text{transmited wave: } \underline{\mathbf{E}}_{t0} &= \underline{\mathbf{E}}_0 (-\sin \theta_2 \mathbf{u}_x + \cos \theta_2 \mathbf{u}_y)\end{aligned}$$

and since according to Eq. (11.19), the tangential component of the electric field (along Oy) is continuous, we have

$$\underline{E}_0 \cos \theta_1 - \underline{E}_{r0} \cos \theta_1 = \underline{E}_{t0} \cos \theta_2, \quad (11.34)$$

or equivalently

$$\cos \theta_1 (1 - r_p) = t_p \cos \theta_2. \quad (11.35)$$



On the other hand, the incident, reflected and transmitted magnetic fields can be obtained from $\underline{\mathbf{B}} = \frac{\mathbf{k}}{\omega} \times \underline{\mathbf{E}}$, yielding:

$$\underline{\mathbf{B}}_{i0} = \underbrace{\frac{n_1}{c}}_{k/\omega} (\cos \theta_1 \mathbf{u}_x + \sin \theta_1 \mathbf{u}_y) \times \underline{\mathbf{E}}_{i0} = \frac{n_1}{c} \underline{\mathbf{E}}_0 \mathbf{u}_z,$$

and similarly

$$\begin{aligned}\underline{\mathbf{B}}_{r0} &= \frac{n_1}{c} \underline{\mathbf{E}}_{r0} \mathbf{u}_z, \\ \underline{\mathbf{B}}_{t0} &= \frac{n_2}{c} \underline{\mathbf{E}}_{t0} \mathbf{u}_z.\end{aligned}$$

Assuming a current-free interface, we can impose the continuity condition (11.20) for the tangential component (along Oz) of the magnetic field:

$$n_1 (\underline{\mathbf{E}}_0 + \underline{\mathbf{E}}_{r0}) = n_2 \underline{\mathbf{E}}_{t0}, \quad (11.36)$$

or equivalently

$$n_1 (1 + r_p) = n_2 t_p, \quad (11.37)$$

Combining Eqs. (11.35) and (11.37) allows us to find Fresnel's coefficients giving the reflected and transmitted amplitudes in terms of the incident one for a p -polarized wave:

$$\begin{aligned}r_p &= \frac{\underline{\mathbf{E}}_{r0}}{\underline{\mathbf{E}}_0} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_1 \cos \theta_2 + n_2 \cos \theta_1}, \\ t_p &= \frac{\underline{\mathbf{E}}_{t0}}{\underline{\mathbf{E}}_0} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_2 + n_2 \cos \theta_1}.\end{aligned} \quad (11.38)$$

11.6.3 Brewster's angle and polarization by reflection

Consider the case of two dielectric media far from any resonance so that propagation is lossless in both media. Remarkably, the reflection coefficient for a p -polarized wave vanish at a specific incident

angle θ_B , called the Brewster angle, such that $r_p = 0$ in Eq. (11.38):

$$\begin{aligned} n_2 \cos \theta_B &= n_1 \cos \theta_t^B = n_1 \sqrt{1 - \sin^2 \theta_t^B} \\ &= n_1 \sqrt{1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2 \theta_B} \end{aligned}$$

and so

$$\left(\frac{n_2}{n_1}\right)^2 \cos^2 \theta_B = 1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2 \theta_B.$$

Finally, since $1 = \sin^2 \theta_B + \cos^2 \theta_B$, the above equation reads

$$\frac{n_2^2 - n_1^2}{n_1^2} \cos^2 \theta_B = \frac{n_2^2 - n_1^2}{n_2^2} \sin^2 \theta_B$$

and its solution is

$$\boxed{\theta_B = \tan^{-1} \left(\frac{n_2}{n_1} \right).} \quad (11.39)$$

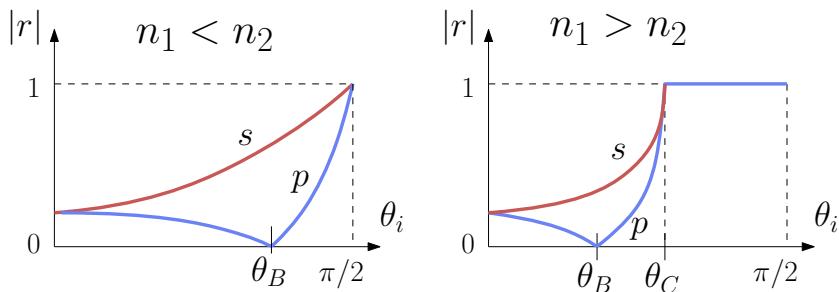
In addition, when the incident angle equals the Brewster angle, the transmitted angle is such that

$$\sin \theta_t^B = \underbrace{\frac{n_1}{n_2} \sin(\tan^{-1} \left(\frac{n_2}{n_1} \right))}_{\theta_B} = \frac{n_1}{n_2} \frac{n_2/n_1}{\sqrt{1 + (n_2/n_1)^2}} = \frac{n_1}{\sqrt{n_1^2 + n_2^2}} = \cos \theta_B$$

and we conclude that

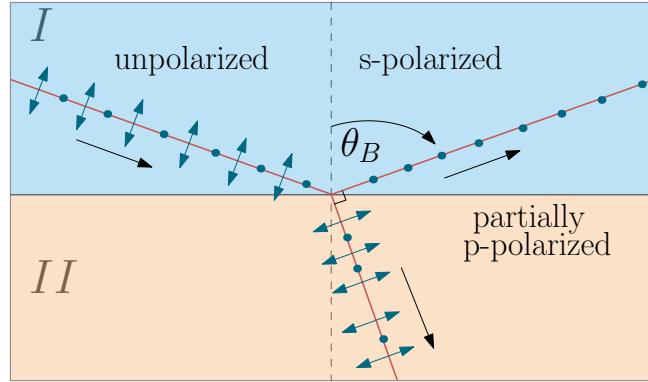
$$\theta_t^B = \frac{\pi}{2} - \theta_B.$$

Note that the Brewster angle has no equivalent in the case of s polarization. Indeed, imposing $r_s = 0$ in Eq. (11.33) would lead to $\tan^2 \theta_B < 0$. The following figure shows the general incident angle dependence of the reflection coefficient at an interface separating two dielectrics, the reflection vanishes at $\theta_i = \theta_B$ for p -polarized light. Note that for $n_1 < n_2$ and in the limit of grazing incidence ($\theta_i \rightarrow \pi/2$), the reflection coefficient tends to 1. Thus, any dielectric interface becomes a good mirror at grazing incidence. In the case $n_1 > n_2$, there is total reflection for $\theta > \theta_C = \sin^{-1} \left(\frac{n_2}{n_1} \right) > \theta_B$.

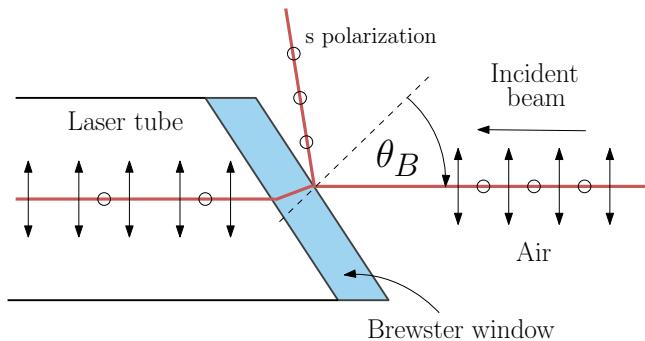


The reflection of light at an air-dielectric interface is another way to obtain polarized light, alternative to the use of transmission polarizers discussed in the course of electromagnetism and light. Indeed, when unpolarized light hits a dielectric interface at the Brewster angle, the reflected wave is completely

s polarized. This phenomenon of polarization by reflection was reported by Malus in 1809 who observed the reflection of sunlight by a glass window at the Brewster angle. Yet, it is Brewster who correctly formulated (in 1815) the dependence of Brewster's angle upon the refractive index ratio across an interface.



In some laser cavities, a window tilted at Brewster's angle with respect to the optical axis (also called the Brewster window) allows for controlling the polarization of the output laser beam. Indeed, an *s*-polarized beam will lose intensity every time it encounters the Brewster window in the cavity, and so the amplified light emission, which makes several roundtrips after exiting the laser tube, will be exclusively *p*-polarized.



11.6.4 Fresnel's coefficients for an arbitrary polarization

In the most general case, the amplitude of the incident electric field can be always written as a linear superposition of *s*- and *p*-polarized fields

$$\underline{\mathbf{E}}_i = \alpha_1 \underline{\mathbf{E}}_s + \alpha_2 \underline{\mathbf{E}}_p \quad \alpha_1, \alpha_2 \in \mathbb{C}$$

for which the reflected and transmitted amplitudes are simply

$$\begin{aligned} \underline{\mathbf{E}}_r &= \alpha_1 r_s \underline{\mathbf{E}}_s + \alpha_2 r_p \underline{\mathbf{E}}_p, \\ \underline{\mathbf{E}}_t &= \alpha_1 t_s \underline{\mathbf{E}}_s + \alpha_2 t_p \underline{\mathbf{E}}_p. \end{aligned}$$

11.6.5 Normal incidence

Note that for the particular case of normal incidence, the Snell-Descartes law gives $\theta_2 = \theta_1 = 0$, so that the reflection and transmission coefficients are equal for both s- and p-polarization:

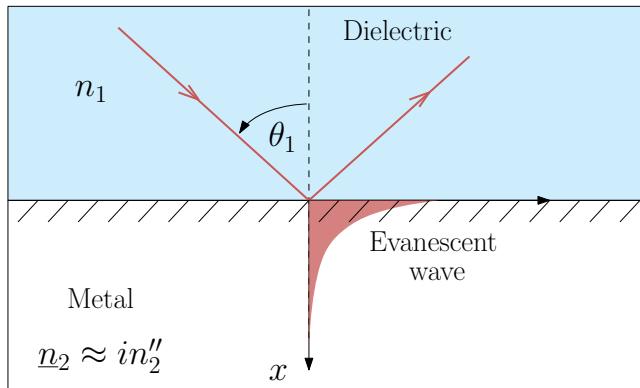
$$\boxed{\begin{aligned} r_p &= \frac{\underline{E}_{r0}}{\underline{E}_0} = \frac{n_2 - n_1}{n_1 + n_2} = -r_s, \\ t_p &= \frac{\underline{E}_{t0}}{\underline{E}_0} = \frac{2n_1}{n_1 + n_2} = t_p. \end{aligned}} \quad (11.40)$$

11.7 Reflection by metals and metallic waveguides

A metal in the visible range is such that the refractive index is essentially a pure imaginary number, $n_2 \approx i\sqrt{(\omega_p/\omega)^2 - 1}$. If an electromagnetic plane wave is incident from a dielectric medium without absorption, $\underline{n}_1 = n_1 \in \mathbb{R}$. From Fresnel's equations (11.33) and (11.38), we see that the wave is completely reflected by the metallic surface regardless of the incident polarization:

$$\underbrace{\left| \frac{n_1 \cos \theta_1 - i n_2'' \cos \theta_2}{n_1 \cos \theta_1 + i n_2'' \cos \theta_2} \right|^2}_{R_s} = 1 = \underbrace{\left| \frac{i n_2'' \cos \theta_1 - n_1 \cos \theta_2}{n_1 \cos \theta_2 + i n_2'' \cos \theta_1} \right|^2}_{R_p}$$

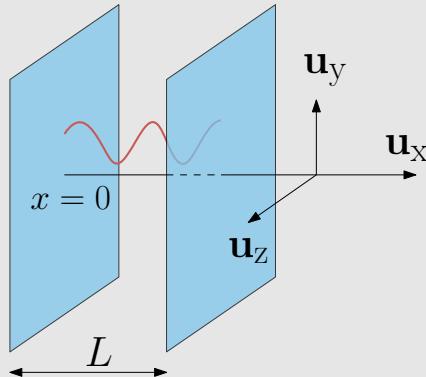
and we conclude from the conservation of energy ($R + 1 = 1$), that $T = 0$. The dielectric-metal interface then constitutes a mirror. In reality, an evanescent wave decays exponentially inside the metal within a typical characteristic length given by the skin depth $\delta = \frac{c}{\omega n_2''}$. In the idealization of a perfect conductor, for which the skin depth δ tends to zero, the electric and magnetic fields are strictly zero inside the metal.



Example 11.3 - Electromagnetic cavity

Consider an electromagnetic wave confined in a cavity made of two perfect conductors as shown in the figure below. The region between the two conductors is delimited by the planes located at $x = 0$ and $x = L$. The electric field of the wave is the following:

$$\underline{\mathbf{E}}(x, t) = \underline{E}_1 e^{i(kx - \omega t)} \mathbf{u}_y + \underline{E}_2 e^{i(-kx - \omega t)} \mathbf{u}_y$$



- a. What is the electric field at $x = 0$ and $x = L$? Deduce \underline{E}_2 as a function of \underline{E}_1 and the possible values k_n that the wavenumber k can take, with $n \in \mathbb{N}^+$.
- b. Determine the relationship between ω and k_n using d'Alembert's wave equation.
- c. Consider a particular mode, that is a particular solution with $\omega = \omega_n$ and $k = k_n$. Write the real electric field $\mathbf{E}_n(x, t)$. How does it evolve in space and time? Does it propagate?
- d. Determine the positions x_p of the nodes, that is, the points for which the electric field is zero for all times t , and give the distance between two successive nodes.
- e. Determine the magnetic field $\mathbf{B}_n(x, t)$ associated with this mode, and determine the magnetic field nodes x'_p .
- f. What is the general form of a one-dimensional (varying with respect to x only) electromagnetic wave in the cavity?

Solution

- a. Inside the conductors, the electromagnetic wave vanish and in particular $\mathbf{E} = \mathbf{0}$. By continuity of the tangential component of the electric field at an interface (see Eq. (11.19)), we have at $x = 0$

$$\underline{\mathbf{E}}(0, t) = \underline{E}_1 e^{-i\omega t} \mathbf{u}_y + \underline{E}_2 e^{-i\omega t} \mathbf{u}_y = \mathbf{0} \quad \forall t$$

and we conclude $\underline{E}_2 = -\underline{E}_1$. Thus, the electric field can be rewritten as

$$\begin{aligned} \underline{\mathbf{E}}(x, t) &= \underline{E}_1 \left(e^{i(kx - \omega t)} - e^{i(-kx - \omega t)} \right) \mathbf{u}_y \\ &= 2i\underline{E}_1 \sin(kx) e^{-i\omega t} \mathbf{u}_y. \end{aligned}$$

The continuity of the tangential component of \mathbf{E} at the interface $x = L$ yields

$$\sin kL = 0 \rightarrow k_n = \frac{n\pi}{L} \quad n \in \mathbb{N}$$

and so

$$\underline{\mathbf{E}}_n(x, t) = 2i\underline{E}_1 \sin(k_n x) e^{-i\omega t} \mathbf{u}_y$$

- b.** The electric field $\underline{\mathbf{E}}_n$ must be a solution to d'Alembert's wave equation. Since it only depends on x

$$\frac{\partial^2 \underline{\mathbf{E}}_n}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \underline{\mathbf{E}}_n}{\partial t^2} = 0.$$

We have

$$\begin{aligned} \frac{\partial^2 \underline{\mathbf{E}}_n}{\partial x^2} &= -2ik_n^2 \underline{E}_1 \cos(k_n x) e^{-i\omega t} \mathbf{u}_y = -k_n^2 \underline{\mathbf{E}}_n, \\ \frac{\partial^2 \underline{\mathbf{E}}_n}{\partial t^2} &= -2i\omega^2 \underline{E}_1 \cos(k_n x) e^{-i\omega t} \mathbf{u}_y = -\omega^2 \underline{\mathbf{E}}_n, \end{aligned}$$

and the introduction of these results in d'Alembert's equation gives the dispersion relation

$$-k_n^2 + \frac{\omega_n^2}{c^2} = 0 \rightarrow \omega_n = ck_n = \frac{n\pi c}{L} \quad n \in \mathbb{N}^+$$

- c.** For the mode n , the real electric field reads

$$\mathbf{E}_n(x, t) = \text{Re}\{2i\underline{E}_1 \sin(k_n x) e^{-i\omega_n t} \mathbf{u}_y\}$$

and by writing $\underline{E}_1 = E_1 e^{i\phi}$, we obtain

$$\mathbf{E}_n(x, t) = 2E_1 \sin(k_n x) \sin(\omega_n t - \phi) \mathbf{u}_y.$$

The wave clearly does not propagate, since it cannot be written as a function of $x \pm ct$. The spatial dependence is always the same, only the overall amplitude oscillates in time. This is called a standing wave.

- d.** The nodes positions x_p are such that

$$\sin k_n x_p = 0 \rightarrow x_p = \frac{p\pi}{k_n} = \frac{pL}{n} = \frac{p\lambda_n}{2},$$

where $\lambda_n = 2\pi/k_n = 2L/n$. Two successive nodes are thus separated by $\lambda_n/2$.

- e.** From Faraday-Maxwell's equation, we have

$$\underbrace{\nabla \times \underline{\mathbf{E}}_n}_{\frac{\partial E_y}{\partial x} \mathbf{u}_z} = -\frac{\partial \mathbf{B}_n}{\partial t}$$

and so $\mathbf{B}_n = B_z \mathbf{u}_z$ with

$$\begin{aligned}\frac{\partial B_z}{\partial t} &= -2E_1 k_n \cos(k_n x) \sin(\omega_n t - \phi) \\ \rightarrow B_z &= \frac{2E_1}{c} \cos(k_n x) \cos(\omega_n t - \phi).\end{aligned}$$

The nodes of the magnetic field are given by $\cos k_n x'_p = 0$, i.e.,

$$x'_p = \frac{(\pi/2 + p\pi)}{k_n} = \frac{\lambda_n}{4} + p \frac{\lambda_n}{2}$$

so that the electric and magnetic field nodes are separated by

$$x'_p - x_p = \frac{\lambda_n}{4}.$$

The most general solution for an electric field depending on x inside the cavity will be a superposition of modes for each of the two possible orthogonal polarizations along \mathbf{e}_y and \mathbf{e}_z

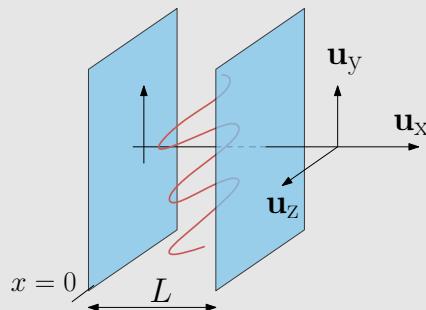
$$\begin{aligned}\mathbf{E}(x, t) &= \sum_n A_n \sin(k_n x) \sin(\omega_n t - \phi_y) \mathbf{u}_y \\ &+ \sum_n B_n \sin(k_n x) \sin(\omega_n t - \phi_z) \mathbf{u}_z.\end{aligned}$$

Example 11.4 - Guided wave

A wave with electric field

$$\mathbf{E} = E_0 \sin \frac{\pi x}{L} \cos(ky - \omega t) \mathbf{u}_z$$

propagates in vacuum between two infinite metallic plates located at $x = 0$ and $x = L$, as shown in the figure below.



- a. What is the propagation direction? Express the wave as a superposition of two plane waves.
- b. What is the value of the electric field on the metallic plates?

- c. Find the conditions under which this wave is a solution to the propagation equation. What are the possible values of the frequency ω that allows for a propagating wave inside the guide?
- d. Express the magnetic field associated with \mathbf{E} .
- e. Calculate the average (in both space and time) volume energy density and the average Poynting vector.
- f. What is the propagation velocity \mathbf{v}_e for the electromagnetic energy?

Solution

- a. The wave is described by a function of $ky - \omega t = k(y - (\omega/k)t)$, so it propagates along the direction of increasing values of y at a speed ω/k . Using the identity

$$\sin A \cos B = \frac{1}{2} (\sin(A - B) + \sin(A + B)),$$

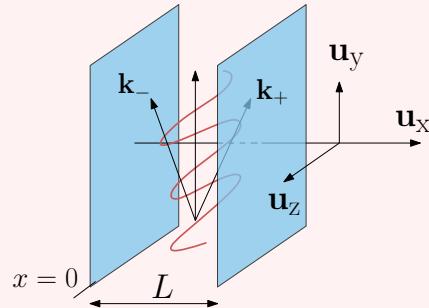
we can rewrite the electric field as

$$\mathbf{E} = \underbrace{\frac{E_0}{2} \sin \left(ky + \frac{\pi x}{L} - \omega t \right) \mathbf{u}_z}_{\mathbf{k}_+ \cdot \mathbf{x} - \omega t} - \underbrace{\frac{E_0}{2} \sin \left(ky - \frac{\pi x}{L} - \omega t \right) \mathbf{u}_z}_{\mathbf{k}_- \cdot \mathbf{x} - \omega t}.$$

We recognize the superposition of two plane waves having the same frequency ω and wavevectors

$$\mathbf{k}_+ = k \mathbf{u}_y + \frac{\pi}{L} \mathbf{u}_x, \quad \mathbf{k}_- = k \mathbf{u}_y - \frac{\pi}{L} \mathbf{u}_x,$$

so that the components of the wavevectors along the x -axis represent counterpropagating waves. The \mathbf{k}_+ (resp. \mathbf{k}_-) wave plays the role of an incident (resp. reflected) wave for the plane wall at $x = L$, and vice-versa for the plane wall at $x = 0$. Each wave propagates between the walls and bounces on them, ensuring under certain condition to be detailed below, that it is guided along the y -direction.



- b. Assuming perfect conductive plates, the wave vanishes everywhere inside them. If no charge is present at the interfaces, the electric field must be continuous (both its normal and tangential components), and so the field is zero at $x = 0$ and $x = L$.

c. The electric field must be a solution to d'Alembert's wave equation

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$

We have

$$\nabla^2 \mathbf{E} = \frac{\partial^2 \mathbf{E}}{\partial x^2} + \frac{\partial^2 \mathbf{E}}{\partial x^y} = - \left(\frac{\pi^2}{L^2} + k^2 \right) \mathbf{E}$$

and $\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\omega^2 \mathbf{E}$ so that the dispersion relation reads:

$$k^2 + \frac{\pi^2}{L^2} = \frac{\omega^2}{c^2}$$

For the wave to propagate along the y -axis, the wavenumber k must be real. This gives a minimum frequency ω for the wave to be able to propagate along the axis of the waveguide

$$k^2 = \frac{\omega^2}{c^2} - \frac{\pi^2}{L^2} > 0$$

and so

$$\omega > \frac{c\pi}{L}.$$

d. The magnetic field can be obtained by using Maxwell-Faraday's law

$$\nabla \times \mathbf{E} = \frac{\partial \mathbf{E}_z}{\partial y} \mathbf{u}_x - \frac{\partial \mathbf{E}_z}{\partial x} \mathbf{u}_y = - \frac{\partial \mathbf{B}}{\partial t}$$

and so

$$\frac{\partial \mathbf{B}}{\partial t} = k E_0 \sin\left(\frac{\pi}{L}x\right) \sin(ky - \omega t) \mathbf{u}_x + \frac{\pi}{L} E_0 \cos\left(\frac{\pi}{L}x\right) \cos(ky - \omega t) \mathbf{u}_y.$$

Integrating with respect to time

$$\mathbf{B} = \frac{k}{\omega} E_0 \sin\left(\frac{\pi}{L}x\right) \cos(ky - \omega t) \mathbf{u}_x - \frac{\pi}{\omega L} E_0 \cos\left(\frac{\pi}{L}x\right) \sin(ky - \omega t) \mathbf{u}_y.$$

e. The electromagnetic energy density u_{EM} writes

$$\begin{aligned} u_{EM} &= \frac{\epsilon_0}{2} |\mathbf{E}|^2 + \frac{1}{2\mu_0} |\mathbf{B}|^2 = \frac{\epsilon_0}{2} (|\mathbf{E}|^2 + c^2 |\mathbf{B}|^2) \\ &= \frac{\epsilon_0}{2} E_0^2 \left(1 + \frac{c^2 k^2}{\omega^2} \right) \sin^2 \left(\frac{\pi x}{L} \right) \cos^2(ky - \omega t) + \frac{\epsilon_0 \pi^2 c^2 E_0^2}{2\omega^2 L^2} \cos^2 \left(\frac{\pi x}{L} \right) \sin^2(ky - \omega t). \end{aligned}$$

The time-averaged density is therefore

$$\langle u_{EM} \rangle_T = \frac{\epsilon_0 E_0^2}{4} \left(1 + \frac{c^2 k^2}{\omega^2} \right) \sin^2 \left(\frac{\pi x}{L} \right) + \frac{\epsilon_0 \pi^2 c^2 E_0^2}{4\omega^2 L^2} \cos^2 \left(\frac{\pi x}{L} \right).$$

Now we can average in the x direction over the length L of the cavity, noting that $\langle \sin^2(\pi x/L) \rangle = \langle \cos^2(\pi x/L) \rangle = 1/2$ and finally

$$\begin{aligned}\langle\langle u_{EM} \rangle_T \rangle &= \frac{\epsilon_0 E_0^2}{8} \left(1 + \underbrace{\frac{c^2 k^2}{\omega^2}}_1 + \underbrace{\frac{\pi^2 c^2}{\omega^2 L^2}}_1 \right) \\ &= \frac{\epsilon_0 E_0^2}{4}.\end{aligned}$$

The Poynting vector is given by

$$\begin{aligned}\boldsymbol{\Pi} &= \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} = \epsilon_0 c^2 \mathbf{E} \times \mathbf{B} \\ &= \epsilon_0 c^2 \frac{E_0^2 k}{\omega} \sin^2 \left(\frac{\pi x}{L} \right) \cos^2(ky - \omega t) \mathbf{u}_y \\ &\quad + \epsilon_0 c^2 \frac{E_0^2 \pi}{2\omega L} \sin \left(\frac{2\pi x}{L} \right) \cos(2(ky - \omega t)) \mathbf{u}_x\end{aligned}$$

whose average over time and space gives

$$\langle\langle \boldsymbol{\Pi} \rangle_T \rangle = \epsilon_0 c^2 \frac{E_0^2 k}{4\omega} \mathbf{u}_y = \frac{c^2 k}{\omega} \langle\langle u_{EM} \rangle_T \rangle \mathbf{u}_y.$$

f. Let S_y be a surface perpendicular to the propagation direction. The flux of the Poynting vector through S_y during a time interval Δt gives an energy flow

$$\Delta E = |\langle\langle \boldsymbol{\Pi} \rangle_T \rangle| S_y \Delta t = v_e \langle\langle u_{EM} \rangle_T \rangle S_y \Delta t,$$

where v_e is the speed at which the energy propagates. Since $|\langle\langle \boldsymbol{\Pi} \rangle_T \rangle| = \frac{c^2 k}{\omega} \langle\langle u_{EM} \rangle_T \rangle$, we obtain

$$\mathbf{v}_e = \frac{c^2 k}{\omega} \mathbf{u}_y = c \sqrt{1 - \left(\frac{\pi c}{\omega L} \right)^2} \mathbf{u}_y.$$

11.8 Reflectance and transmittance

The Fresnel coefficients allow us to determine the amplitude of the reflected and transmitted electric fields at an interface relative to the incident amplitude. If we are interested in the flow of energy, the relevant parameters are the reflectance R and transmittance T defined as the fraction of the electromagnetic energy flux that is reflected and transmitted by the interface, respectively. For this,

we can calculate the incident, reflected and transmitted complex Poynting vectors

$$\begin{aligned} \text{incident wave: } & \underline{\Pi}_i = \frac{\underline{\mathbf{E}}_i \times \underline{\mathbf{B}}_i^*}{2\mu_0} = \frac{|\underline{E}_0|^2}{2\mu_0\omega} \mathbf{k}_i, \\ \text{reflected wave: } & \underline{\Pi}_r = \frac{\underline{\mathbf{E}}_r \times \underline{\mathbf{B}}_r^*}{2\mu_0\omega} = \frac{|r \underline{E}_0|^2}{2\mu_0\omega} \mathbf{k}_r, \\ \text{transmited wave: } & \underline{\Pi}_t = \frac{\underline{\mathbf{E}}_t \times \underline{\mathbf{B}}_t^*}{2\mu_0} = \frac{|t \underline{E}_0|^2}{2\mu_0\omega} \mathbf{k}_t, \end{aligned}$$

and recalling that the average power propagating along the normal to the interface (irradiance) is given by the projection of $\langle \underline{\Pi} \rangle_T = \operatorname{Re}\{\underline{\Pi}\}$ along the normal (here \mathbf{u}_x), we find

$$\begin{aligned} \text{incident: } & \mathcal{I}_i = \operatorname{Re}\{\underline{\Pi}_i\} \cdot \mathbf{u}_x = \frac{|\underline{E}_0|^2}{2\mu_0\omega} k_1 \cos \theta_1, \\ \text{reflected: } & \mathcal{I}_r = \operatorname{Re}\{\underline{\Pi}_r\} \cdot (-\mathbf{u}_x) = \frac{|r \underline{E}_0|^2}{2\mu_0\omega} k_1 \cos \theta_1, \\ \text{transmited: } & \mathcal{I}_t = \operatorname{Re}\{\underline{\Pi}_t\} \cdot \mathbf{u}_x = \frac{|t \underline{E}_0|^2}{2\mu_0\omega} k_2 \cos \theta_2, \end{aligned}$$

and since $k_1 = \frac{\omega}{c} n_1$, $k_2 = \frac{\omega}{c} n_2$, we find the reflectance and transmittance at an interface in terms of the Fresnel coefficients:

$$R = \frac{\mathcal{I}_r}{\mathcal{I}_i} = r^2,$$

$$T = \frac{\mathcal{I}_t}{\mathcal{I}_i} = t^2 \frac{n_2 \cos \theta_2}{n_1 \cos \theta_1}.$$

(11.41)

We note that $R + T = 1$ in both cases of s - and p -polarizations, in keeping with the conservation of energy: the incident power flux reflected with a fraction R and transmitted across the interface with a fraction T .

11.9 Appendix: Dispersion, temporal response and causality

11.9.1 Dispersion

A medium is dispersive if its response depends on frequency. The cause for dispersion is quite general. It comes from the non-instantaneous response of the medium to an external excitation. That is the case for the polarization \mathbf{P} of a dielectric medium under the influence of an electromagnetic wave (or the conduction current in the case of a conductor). Whereas in the static regime, one may write $\mathbf{P} = \epsilon_0 \chi \mathbf{E}$, where χ is the susceptibility, in the time-dependent regime we generally have

$$\mathbf{P}(\mathbf{x}, t) \neq \epsilon_0 \chi \mathbf{E}(\mathbf{x}, t),$$

since friction and inertia are responsible for a retarded response of the medium polarization to the varying electric field. The polarization vector at instant t depends not only on the value of the electric field at t but also on the history of the electric field, that is, on the previous values of the field at times preceding t . For a linear, homogeneous and isotropic medium, the polarization generally reads

$$\mathbf{P}(\mathbf{x}, t) = \epsilon_0 \int_{-\infty}^{\infty} \chi(t') \mathbf{E}(\mathbf{x}, t - t') dt' \quad (11.42)$$

we recognize in Eq. (11.42) a convolution product which makes dealing with Maxwell's equations difficult a priori. On the other hand, \mathbf{P} can always be written as a superposition of sinusoidal waves due to Fourier's theorem

$$\mathbf{P}(\mathbf{x}, t) = \int_{-\infty}^{\infty} \underline{\mathbf{P}}(\mathbf{x}, \omega) e^{i\omega t} dt$$

where $\underline{\mathbf{P}}(\mathbf{x}, \omega)$ is the Fourier transform of \mathbf{P} with respect to time. Equation (11.42) and the properties of the Fourier transform then implies that

$$\underline{\mathbf{P}}(\mathbf{x}, \omega) = \epsilon_0 \underline{\chi}(\omega) \underline{\mathbf{E}}(\mathbf{x}, \omega),$$

where the complex susceptibility $\underline{\chi}(\omega)$ is the Fourier transform of χ . This last equality and the linearity of Maxwell's equations justifies working with monochromatic plane waves: the response of the medium to a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}(\omega) e^{i(\underline{k}x - \omega t)}$ simply reads

$$\underline{\mathbf{P}} = (\underline{\epsilon}(\omega) - \epsilon_0) \underline{\mathbf{E}} = \epsilon_0 \underline{\chi}(\omega) \underline{\mathbf{E}}$$

11.9.2 Causality

This section presents general properties of the susceptibility that are a consequence of the causality principle. Namely, relations called Kramers-Kronig relations are established between the real and imaginary parts of the complex susceptibility. Complex analysis, integration in the complex plane and mathematical notions about distributions constitute prerequisites to follow the detailed derivation of Kramers-Kronig relations, which can thus be skipped in a first reading.

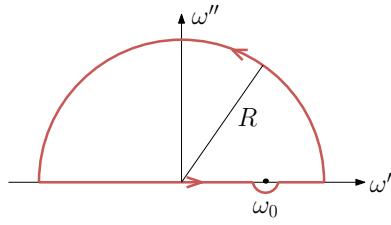
Causality imposes that in Eq. (11.42), $\chi(t) = 0$ if $t < 0$. Otherwise, the polarization at time t would depend on future values of the electric field, violating the causality principle that states that any effect cannot precede its cause. We will now show that causality introduces a constraint linking the real and imaginary parts of the susceptibility $\underline{\chi}(\omega)$ (and therefore, of the complex permittivity). Fourier's inversion formula together with $\chi = 0$ for $t < 0$ implies

$$\underline{\chi}(\omega) = \frac{1}{2\pi} \int_0^{\infty} \chi(t) e^{+i\omega t} dt.$$

To calculate the latter integral, we may use an extension of the complex susceptibility $\underline{\chi}$ to the complex frequency plane, that is, we consider $\omega = \omega' + i\omega''$, and note that for $\text{Im}\{\omega\} = \omega'' > 0$, we have

$$|\underline{\chi}(\omega)| < \frac{1}{2\pi} \int_0^{\infty} |\chi(t) e^{+i\omega t}| dt < \frac{1}{2\pi} \int_0^{\infty} |\chi(t)| dt,$$

so that $\underline{\chi}(\omega)$ extended to the complex plane is bounded for every ω such that $\text{Im}\{\omega\} > 0$ and therefore has no poles in this half plane. The function $\underline{\chi}(\omega)$ is therefore analytic. We now define the function $f(\omega) = \underline{\chi}(\omega)/(\omega - \omega_0)$ where ω_0 is real. This function has a pole in ω_0 and we can apply Cauchy's residues theorem by integrating f over the closed curve shown in the following figure.



The integral over the half-circle tends to zero when $R \rightarrow \infty$ provided that $\lim_{|\omega| \rightarrow \infty} |\chi(\omega)| = 0$, which is the case since all materials become transparent at high frequencies due to the inertia of charged particles. Cauchy's theorem then gives

$$i\pi\underline{\chi}(\omega_0) + P \int \frac{\underline{\chi}(\omega')}{\omega' - \omega_0} d\omega' = 2i\pi\underline{\chi}(\omega_0)$$

where P stands for the principal part of the integral. We obtain

$$\underline{\chi}(\omega_0) = \frac{1}{i\pi} P \int_{-\infty}^{\infty} \frac{\underline{\chi}(\omega')}{\omega' - \omega_0} d\omega'$$

and by taking the real and imaginary part and by writing $\underline{\chi}(\omega) = \underline{\epsilon}(\omega)/\epsilon_0 - 1$, we obtain the Kramers-Kronig relations,

$$\begin{aligned} \frac{\epsilon_R(\omega_0)}{\epsilon_0} &= 1 + \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\epsilon_I(\omega')/\epsilon_0}{\omega' - \omega_0} d\omega', \\ \frac{\epsilon_I(\omega_0)}{\epsilon_0} &= -\frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\epsilon_R(\omega')/\epsilon_0 - 1}{\omega' - \omega_0} d\omega'. \end{aligned} \quad (11.43)$$

Considering finally that $\epsilon(t)$ is real, we have $\underline{\epsilon}(-\omega) = \underline{\epsilon}^*(\omega)$ and

$$\frac{\epsilon_R(\omega_0)}{\epsilon_0} = 1 + \frac{2}{\pi} P \int_{-\infty}^{\infty} \frac{\omega' \epsilon_I(\omega')/\epsilon_0}{\omega'^2 - \omega_0^2} d\omega'. \quad (11.44)$$

This relation allows us, for example, to determine the real part of the permittivity from absorption measurements.

Remark

The Kramers-Kronig relations may be more easily obtained by using some well known results of Fourier analysis. Indeed, since causality forces the function χ to be zero for $t < 0$, this means that χ can be written as

$$\chi(t) = \chi(t)\Theta(t),$$

where Θ is the Heaviside function, that is,

$$\Theta(t) = \begin{cases} 1 & \text{if } t > 0, \\ 0 & \text{if } t < 0. \end{cases}$$

It follows that

$$\underline{\chi}(\omega) = (\underline{\Theta} * \underline{\chi})(\omega),$$

where $*$ stands for convolution. The Fourier transform of the Heaviside (in the framework of distributions) reads

$$\underline{\Theta}(\omega) = \frac{1}{2\pi i} \text{p.v.} \frac{1}{\omega} + \frac{1}{2}\delta(\omega),$$

where p.v stands for Cauchy's principal value. Finally,

$$\chi(\omega) = \frac{1}{2}\chi(\omega) - \frac{i}{2\pi}P \int_{-\infty}^{\infty} \frac{\chi(\omega')}{\omega - \omega'} d\omega',$$

which leads to the Kramers-Kronig relations.

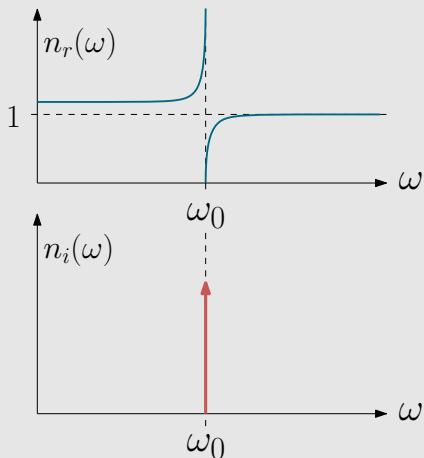
Example 11.5 - Absorption peak and permittivity

A very narrow absorption peak at frequency ω_0 may be modeled by a Dirac distribution of the form

$$\epsilon_I(\omega) = f\delta(\omega - \omega_0).$$

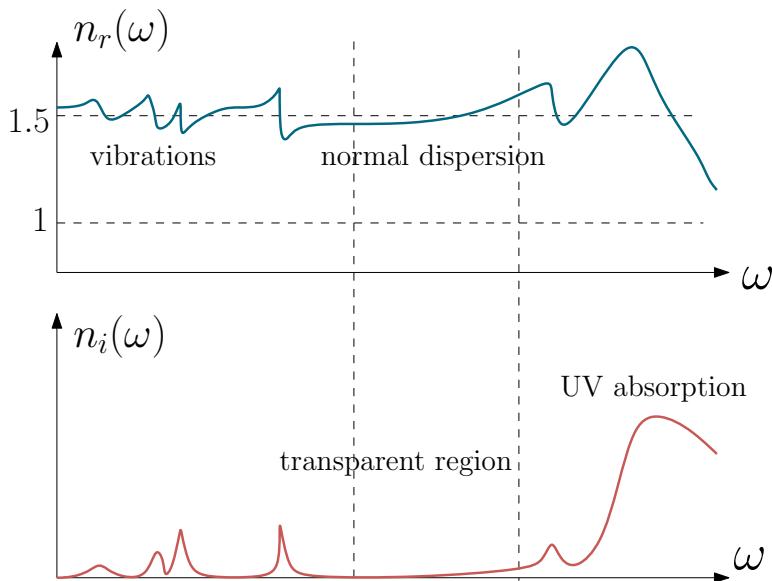
Kramers-Kronig's relation (11.44) then states that this contributes to dispersion, giving a real part of the permittivity of the form

$$\epsilon_R(\omega) = \epsilon_0 \left(1 + \frac{2}{\pi\epsilon_0} \int_{\mathbb{R}} \frac{f\omega' \delta(\omega' - \omega_0)}{\omega'^2 - \omega^2} d\omega' \right) = \epsilon_0 \left(1 + \frac{2\omega_0}{\pi\epsilon_0} \frac{f}{\omega_0^2 - \omega^2} \right).$$



11.9.3 Normal dispersion, Cauchy's law

Typical dielectrics used in optics (glass, quartz, plastic, silica) have an index of refraction that is quite transparent in the visible range and exhibits a so-called *normal* dispersion, that is, an index of refraction that increases at shorter wavelengths (higher frequencies), as shown in the figure below



Indeed, away and below the resonance ($\omega_0^2 - \omega^2 \gg \gamma\omega$) where there is no absorption, Eq. (11.16) can be approximated by

$$n(\omega) \approx 1 + \frac{1}{2} \frac{\omega_p^2}{(\omega_0^2 - \omega^2)} \approx 1 + \frac{1}{2} \frac{\omega_p^2}{\omega_0^2} \left(1 + \frac{\omega^2}{\omega_0^2} \right) = n_R(\omega). \quad (11.45)$$

In terms of the wavelength in vacuum, $\lambda = 2\pi c/\omega$, we obtain Cauchy's law,

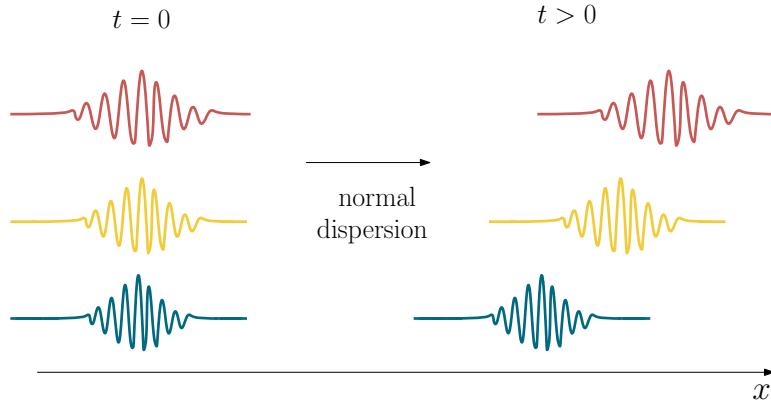
$$n(\lambda) \approx A + \frac{B}{\lambda^2}, \quad (11.46)$$

with $A = 1 + 1 + \frac{1}{2} \frac{\omega_p^2}{\omega_0^2}$ and $B = \frac{2\pi^2 \omega_p^2 c^2}{\omega_0^4}$, which explains the dispersion of sunlight by a prism. Normal dispersion implies $n(\text{blue}) > n(\text{red})$ so that blue light is more deviated than red.

11.10 Appendix: Dispersion and propagation of wave packets

This section illustrates how the general results on the propagation of sinusoidal plane waves in dielectrics allows us to construct solutions to Maxwell's equations that are localized in time and space. This is a bridge to the optics course which can be seen as a reminder.

We have seen that, quite generally, the refractive index of a medium depends on frequency. This is at the origin of dispersion: a light pulse having different frequency components will be deformed during propagation, since different wavelengths travel at different speeds in matter. From the Lorentz model, we see that below a resonance $n_R(\omega)$ increases with ω , a regime called normal dispersion and that represents the typical behavior of transparent materials used in optics (illustrated in the figure below). In contrast, above a resonance we see that $n_R(\omega)$ decreases with ω , a regime called *anomalous* dispersion.



11.10.1 Wave packets

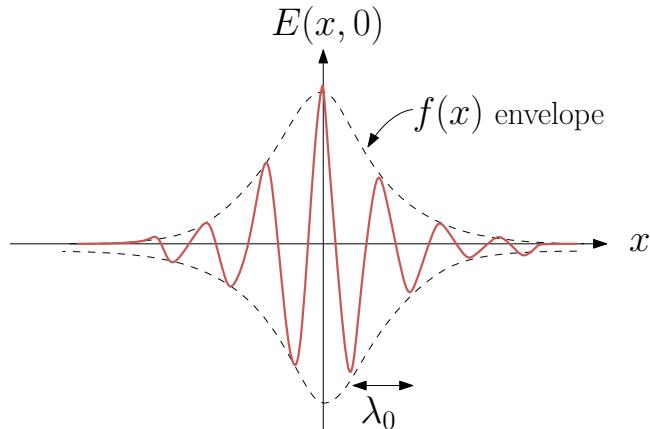
Consider, for simplicity, a plane-wave of frequency ω_0 propagating along the x -axis and polarized along \mathbf{u}_y

$$\mathbf{E}(x, t) = E_0 \cos(k(\omega_0)x - \omega_0 t) \mathbf{u}_y, \quad (11.47)$$

where the dispersion relation in a dielectric far from any absorption reads $k(\omega) = n(\omega)w/c$. Of course, such a wave cannot represent a real light pulse since it has an infinite extension in space and time. A way to build a realistic wave with a finite temporal and spatial extent, and centered around a main frequency ω_0 would be to multiply the plane wave (11.47) by an envelope function $f(x)$ peaked around $x = 0$. In this way, at $t = 0$ the amplitude of the electric field reads

$$E(x, t = 0) = E_0 f(x) \cos(k(\omega_0)x),$$

which has a finite extension around $x = 0$ at $t = 0$, as shown in the figure below.



Remarkably, as long as $E(x, 0)$ is an integrable function, it can always be written as a linear superposition of plane waves e^{ikx} according to Fourier's theorem,

$$E(x, 0) = \int_{\mathbb{R}} \underline{E}(k) e^{ikx} dk,$$

where $\underline{E}(k)$ is a continuous, complex-valued function of k that decays to zero at infinity, called the Fourier transform of $E(x, 0)$. In order to obtain a wave packet $E(x, t)$ that satisfies the propagation equation in the medium, all we need to do is to shift each plane wave e^{ikx} by its corresponding distance traveled during a time t , that is,

$$\underline{E}(x, t) = \int_{\mathbb{R}} \underline{E}(k) \underbrace{e^{i(kx - \omega(k)t)}}_{\text{plane wave}} dk. \quad (11.48)$$

In this way, it is clear that $E(x, t)$ is a linear superposition of sinusoidal plane waves, which, according to the superposition principle, is a solution to the wave equation.

11.10.2 Evolution of a wave packet in the linear regime

Consider first the case of a wave packet made of plane waves having wave numbers very close to a central wave number k_0 . This is equivalent to ask for the Fourier transform $\underline{E}(k)$ to be peaked around $k = k_0$. In this case, the variations of $n(k)$ and therefore $\omega(k)$ are not too large, and we can write up to first order in $k - k_0$,

$$\omega(k) \approx \omega(k_0) + \left(\frac{\partial \omega}{\partial k} \right)_{k_0} (k - k_0) = \omega_0 + v_g(k - k_0), \quad (11.49)$$

where $\omega_0 = \omega(k_0)$ is the central frequency of the wave packet and $v_g = \left(\frac{\partial \omega}{\partial k} \right)_{k_0}$ is called the group velocity. This linear approximation corresponds to the case of a constant refractive index n . Replacing this expansion in Eq. (11.48), we obtain

$$\underline{E}(x, t) = \underbrace{e^{i(k_0 x - \omega_0 t)}}_{\text{carrier at frequency } \omega_0} \underbrace{\int_{\mathbb{R}} \underline{E}(k) e^{i(k - k_0)(x - v_g t)} dk}_{\text{envelope } f(x - v_g t)}.$$

The wave packet is, at every time, the product of a sinusoidal plane wave at frequency ω_0 modulated by an envelope f that propagates without deformation with velocity equal to the group velocity v_g . Thus,

$$|f(x - v_g t)| = E(x - v_g t, 0).$$

Definition 11.1: Phase and group velocity

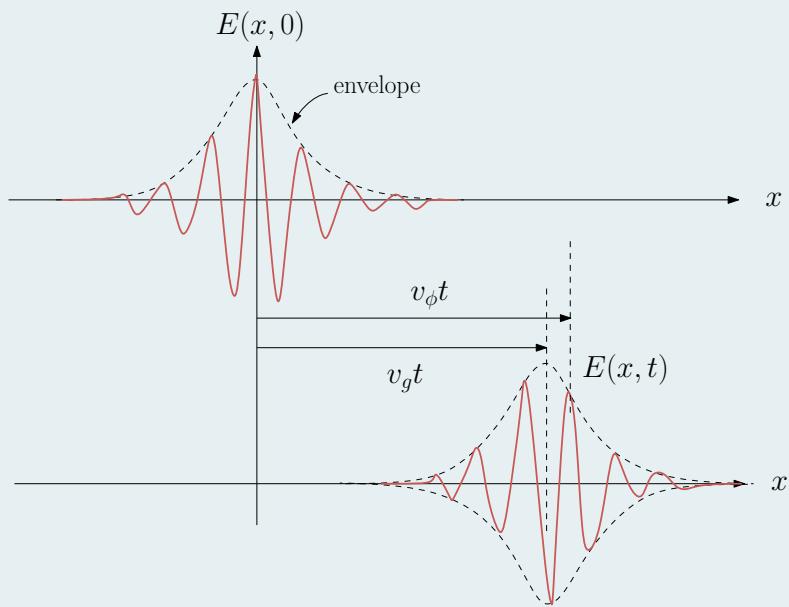
The phase velocity v_ϕ is defined as

$$v_\phi = \frac{\omega}{k(\omega)} = \frac{c}{n(\omega)} \quad (11.50)$$

and the group velocity v_g as

$$v_g = \frac{\partial \omega}{\partial k}. \quad (11.51)$$

For a wave packet composed of plane waves centered at frequency ω_0 , the phase velocity v_ϕ represents the speed of the crests of the wave, whereas the group velocity v_g is the speed of the envelope, as shown in the figure below.



Note that the velocity of the envelope is the velocity at which the intensity travels. Writing $\omega = ck/n$, we have

$$v_g = \frac{c}{n} + ck \frac{\partial}{\partial k} \left(\frac{1}{n} \right) = \underbrace{\frac{c}{n}}_{v_\phi} + \underbrace{ck \frac{\partial}{\partial \omega} \left(\frac{1}{n} \right)}_{= -\frac{\omega}{n} \frac{\partial n}{\partial \omega}} \underbrace{\frac{\partial \omega}{\partial k}}_{v_g}$$

and so the phase and group velocities are related by

$$v_g = \frac{v_\phi}{1 + \frac{\omega}{n} \frac{\partial n}{\partial \omega}} = \frac{c/n}{1 + \frac{\omega}{n} \frac{\partial n}{\partial \omega}}.$$

(11.52)

11.10.3 Dispersion and deformation of wave packets

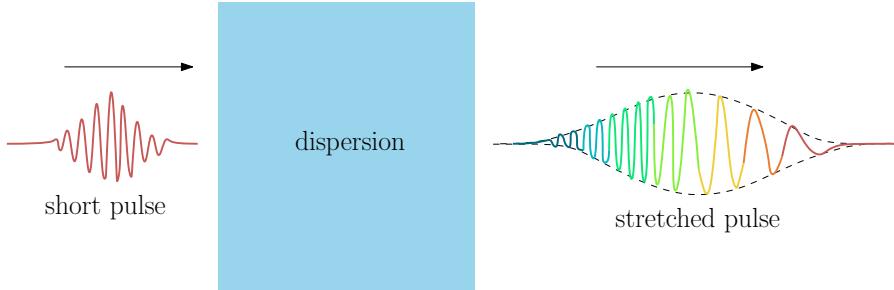
When the dispersion relation $\omega(k)$ is not linear, which is the case for electromagnetic waves in matter, the linear approximation (11.49) will cease to be valid after the wave packet has propagated for a certain time. Eventually, the frequency dependence of the optical index will lead to a deformation of the wave packet. The next term in the expansion would lead to

$$\omega(k) = \omega_0 + \left(\frac{\partial \omega}{\partial k} \right)_{k_0} (k - k_0) + \frac{1}{2} \left(\frac{\partial^2 \omega}{\partial k^2} \right)_{k_0} (k - k_0)^2 + \dots,$$

which can be rewritten as

$$\omega(k) \approx \omega_0 + \underbrace{\left\{ \left(\frac{\partial \omega}{\partial k} \right)_{k_0} + \frac{1}{2} \left(\frac{\partial^2 \omega}{\partial k^2} \right)_{k_0} (k - k_0) \right\}}_{v_g^{\text{eff}}} (k - k_0),$$

where v_g^{eff} is an effective group velocity that depends on wavelength. This leads to a deformation of the wave packet: for example after passing through a material with normal dispersion, red wavelengths will accumulate at the front, leaving the blue components at the back. This stretched light pulse gets therefore longer in time and it has a frequency that depends on time, which is called a *frequency chirp*.



11.10.4 An application of dispersion that led to a Nobel prize

Intense and ultrashort laser pulses, with duration in the femtosecond range, are difficult to implement since their extremely high peak intensity may damage optical components and specially amplifier crystals. An elegant method to solve this issue is to first stretch the pulse, thanks to a dispersive element, before going into the amplifier. The pulse is thus stretched from the femtosecond to the nanosecond range, such that high frequencies pass through the amplifier before the lower frequencies (such a pulse is called *chirped*). The intensity is therefore spread over a longer time and the peak power is reduced by several orders of magnitude, excluding non linear effects and possible damage during the amplification stage. Finally, the pulse is compressed back by inverting the initial dispersion. The physicists who implemented this *chirped pulse amplification*, G. Mourou and D. Strickland, were awarded the Physics Nobel prize in 2018.

11.11 Summary and essential formulas

- In linear, homogeneous and isotropic (LHI) media and in the presence of conduction charges of density $\rho(\mathbf{r}, t)$ and currents of density $\mathbf{J}(\mathbf{r}, t)$, the electric and magnetic fields satisfy a non-homogeneous d'Alembert wave equation

$$\nabla^2 \mathbf{E} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \mathbf{E} = \frac{\nabla \rho}{\epsilon} + \mu \frac{\partial \mathbf{J}}{\partial t},$$

$$\nabla^2 \mathbf{B} - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \mathbf{B} = -\mu \nabla \times \mathbf{J},$$

where $v = 1/\sqrt{\epsilon\mu}$ is the speed of light in the medium, ϵ , the permittivity of the medium and μ its permeability.

The case of nonmagnetic materials is retrieved by setting $\mu = \mu_0$ and the case of metals by using the permittivity and permeability of vacuum, $\epsilon = \epsilon_0$ and $\mu = \mu_0$.

- Plane waves propagating in a LHI medium without absorption have the same structure as in vacuum, that is, if \mathbf{n} denotes the unit vector along the propagation direction, the fields \mathbf{E} and \mathbf{B} are transverse and $(\mathbf{E}, \mathbf{B}, \mathbf{n})$ form a right-handed orthogonal triad satisfying

$$\mathbf{B} = \frac{\mathbf{n}}{v} \times \mathbf{E}.$$

In addition the components of the electric and magnetic field satisfy a homogeneous d'Alembert equation and can be expressed quite generally as the superposition of a forward and a backward-propagating plane wave. For instance, if the propagation direction is the x -axis, the E_y component of the electric field reads

$$E_y = \underbrace{f_y(x - vt)}_{\text{forward}} + \underbrace{g_y(x + vt)}_{\text{backward}},$$

where f and g are arbitrary functions and *forward* or *backward* refers to increasing or decreasing values of x , respectively.

- Sinusoidal plane waves are represented in complex notation by complex fields $\underline{\mathbf{E}}$ such that the physical quantity under consideration is the real part of the complex quantity: $\mathbf{E} = \operatorname{Re}(\underline{\mathbf{E}})$. For a wave of frequency $\omega \in \mathbb{R}$,

$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{-i\omega t + i\mathbf{k} \cdot \mathbf{r}},$$

where the amplitude $\underline{\mathbf{E}}_0$ is a vector with complex valued components and the wavevector \mathbf{k} as well, allowing us to deal with media with absorption.

Faraday-Maxwell's equation lead to the relation involving the right-handed prthogonal triad $(\underline{\mathbf{E}}, \underline{\mathbf{B}}, \mathbf{k})$:

$$\underline{\mathbf{B}} = \frac{\mathbf{k}}{\omega} \times \underline{\mathbf{E}}.$$

- Due to the inertia of charges in matter, the response to an oscillating electric field is generally not instantaneous. This leads to a frequency-dependence of the permittivity, and of the conductivity of metals, which is at the origin of dispersion. These quantities are derived from the matter response to a sinusoidal plane wave, and used in constitutive relations of a medium using complex notations. For a dielectric (LHI medium):

$$\underline{\mathbf{D}}(\mathbf{r}, \omega) = \underline{\epsilon}(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega),$$

and for a metal:

$$\underline{\mathbf{J}}(\mathbf{r}, \omega) = \underline{\sigma}(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega).$$

- Sinusoidal plane wave solutions to the non-homogeneous d'Alembert wave equation or to the d'Alembert wave equation with frequency dependent matter response must satisfy a dispersion relation between the real frequency ω and the complex wave vector

$$\underline{\mathbf{k}}(\omega) = k_R(\omega) + ik_I(\omega).$$

The imaginary part of the wavenumber, k_I , is related to the absorption of the electromagnetic energy as the wave propagates in the medium. For a wave propagating along the x -axis, we have

$$\underline{\mathbf{E}}(\mathbf{x}, t) = \underline{\mathbf{E}}_0 e^{i(k_R x - \omega t)} e^{-k_I x},$$

that is an evanescent wave with a typical absorption length $1/k_I$.

- In a conductor, the response to an oscillating electromagnetic wave is characterized by the complex conductivity which, in the Drude-Lorentz model reads

$$\underline{\sigma}(\omega) = \frac{\sigma_0}{1 - i\omega\tau},$$

where σ_0 is the static conductivity and τ the time between two collisions.

- The Lorentz model predicts the frequency dependence of the complex permittivity of dielectrics by modeling an atom as a damped harmonic oscillator with natural frequency ω_0 and damping constant γ . The polarization density $\underline{\mathbf{P}}$ in the presence of a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{i(kx - \omega t)}$ reads

$$\underline{\mathbf{P}} = (\underline{\epsilon}(\omega) - \epsilon_0) \underline{\mathbf{E}},$$

where the complex permittivity is given by

$$\underline{\epsilon}(\omega) = \epsilon_0 \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \right),$$

with $\omega_p = \sqrt{n e^2 / \epsilon_0 m_e}$ the plasma frequency which depends on the density n in the medium and fundamental constants.

- **Dispersion relations**

The dispersion relation for a metal of conductivity $\underline{\sigma}(\omega)$ reads

$$\underline{k}^2 = \frac{\omega^2}{c^2} + i\mu_0\omega\underline{\sigma}(\omega).$$

For a LHI medium of complex permittivity $\underline{\epsilon}(\omega)$, the dispersion relation reads

$$\underline{k}^2(\omega) = \frac{\omega^2}{c^2} \frac{\underline{\epsilon}(\omega)}{\epsilon_0},$$

where c denotes the velocity of light in vacuum. This leads to the definition of a complex refractive index

$$\underline{n}(\omega) = \sqrt{\underline{\epsilon}(\omega)/\epsilon_0}.$$

Dispersion relations of metals and dielectrics show that it is possible to characterize a metal by an effective complex permittivity $\underline{\epsilon}$. Conversely, a dielectric can be characterized by an effective conductivity $\underline{\sigma}$. They are both related by

$$\underline{\epsilon}(\omega) = \epsilon_0 \left(1 + i \frac{\underline{\sigma}(\omega)}{\epsilon_0 \omega} \right).$$

- The phase velocity v_ϕ of a sinusoidal plane wave of frequency ω is defined as

$$v_\phi = \frac{\omega}{k(\omega)} = \frac{c}{n(\omega)}$$

and represents the velocity of the crests of a sinusoidal plane wave of frequency ω . It can be larger than c close to resonances.

- A dispersion relation allows us to obtain the *group velocity* v_g as

$$v_g(\omega_0) = \left. \frac{\partial \omega}{\partial k} \right|_{\omega_0},$$

which represents the propagation velocity of the energy of a wavepacket centered around ω_0 . For a wave packet composed of sinusoidal plane waves centered at frequency ω_0 , the phase velocity v_ϕ represents the speed of the crests of the wave, whereas the group velocity v_g is the speed of the envelope. Both are related by

$$v_g = \frac{v_\phi}{1 + \frac{\omega}{n} \frac{\partial n}{\partial \omega}} = \frac{c/n}{1 + \frac{\omega}{n} \frac{\partial n}{\partial \omega}}.$$

- The real part of the complex Poynting vector in non-magnetized matter

$$\underline{\Pi} = \frac{1}{2\mu_0} \underline{\mathbf{E}} \times \underline{\mathbf{B}}^* \quad (11.53)$$

reads for a sinusoidal plane wave $\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{-i\omega t + i\mathbf{k} \cdot \mathbf{r}}$, propagating along the x -axis:

$$\text{Re}(\underline{\Pi}) = \frac{|\underline{\mathbf{E}}_0|^2}{2\mu_0\omega} k_R e^{-2k_I x} = \mathcal{I}_0 e^{-2k_I x},$$

where

$$\mathcal{I}_0 = \frac{|\underline{\mathbf{E}}_0|^2}{2\mu_0\omega} k_R = \frac{\epsilon_0 c n_R |\underline{\mathbf{E}}_0|^2}{2},$$

is the irradiance at $x = 0$ in the medium and n_R denotes the real part of the refractive index at frequency ω .

- When an electromagnetic wave is incident on an interface separating two dielectric media without absorption, it splits into a reflected and a transmitted (refracted) wave. If \mathbf{u}_x is the normal to the interface, and \mathbf{k}_i the wavevector of an incident plane wave, then the plane of incidence is defined as the plane containing both \mathbf{u}_x and \mathbf{k}_i .

By imposing the continuity of the tangential component of the electric field at the interface, one concludes that the reflected \mathbf{k}_r and transmitted \mathbf{k}_t wavevectors lie also in the plane of incidence. Moreover, by writing $\mathbf{k}_i \cdot \mathbf{u}_x = k_1 \cos \theta_i$, $\mathbf{k}_r \cdot \mathbf{u}_x = -k_1 \cos \theta_r$, and $\mathbf{k}_t \cdot \mathbf{u}_x = k_2 \cos \theta_t$, the continuity condition gives both the reflection law

$$\theta_i = \theta_r$$

and the Snell-Descartes law

$$n_1 \sin \theta_i = n_2 \sin \theta_t$$

where n_1 is the refractive index of the medium from which the wave is incident and n_2 is the refractive index of the medium on which light is refracted.

- When $n_1 > n_2 \in \mathbb{R}$, there is total reflection for angles of incidence above a critical value

$$\theta_i^* = \sin^{-1} \left(\frac{n_2}{n_1} \right).$$

In this case the transmitted wave only propagates along the interface whereas its amplitude decays exponentially in the direction perpendicular to the interface (evanescent wave). An optical fiber operates according to this principle; light is confined inside the fiber core since the fields are totally reflected on its boundary.

- If \underline{E}_0 is the amplitude of the incident electric field, the Fresnel reflection r and transmission t coefficients allow us to express the reflected (\underline{E}_r) and transmitted (\underline{E}_t) amplitudes of the electric field at an interface as

$$\underline{E}_{r0} = r \underline{E}_0,$$

$$\underline{E}_{t0} = t \underline{E}_0.$$

For an *s*-polarized wave (electric field perpendicular to the plane of incidence), the Fresnel coefficients are

$$r_s = \frac{\underline{E}_{r0}}{\underline{E}_0} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2},$$

$$t_s = \frac{\underline{E}_{t0}}{\underline{E}_0} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + n_2 \cos \theta_2},$$

whereas for a *p*-wave (electric field in the plane of incidence), they are

$$r_p = \frac{\underline{E}_r}{\underline{E}_0} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_1 \cos \theta_2 + n_2 \cos \theta_1},$$

$$t_p = \frac{\underline{E}_t}{\underline{E}_0} = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_2 + n_2 \cos \theta_1}.$$

At normal incidence, we have

$$r_p = -r_s = \frac{n_2 - n_1}{n_2 + n_1}$$

and

$$t_s = t_p = \frac{2n_1}{n_1 + n_2}.$$

- For a *p*-polarized wave, there exists a particular incident angle, called the Brewster angle θ_B for which there is no reflection at the interface ($r_p = 0$). If an initially unpolarized beam of light strikes an interface at Brewster's angle, then the reflected wave is perfectly *s*-polarized and the transmitted wave partially *p*-polarized. Such an interface acts as a linear polarizer. Brewster's windows are used in laser cavities which transmit the *p*-polarized component of the electric field but introduce losses for the *s*-polarized component. The laser beam at the output is therefore *p*-polarized.
- The reflectance R and transmittance T represent the fraction of the incident electromagnetic energy flux that is reflected by and transmitted through the interface, respectively. They are related to Fresnel's reflection and transmission coefficients by

$$R = \frac{\mathcal{I}_r}{\mathcal{I}_i} = r^2,$$

$$T = \frac{\mathcal{I}_t}{\mathcal{I}_i} = t^2 \frac{n_2 \cos \theta_2}{n_1 \cos \theta_1}.$$

Conservation of the energy flux at an interface in the absence of absorption yields

$$R + T = 1.$$

Chapter 12

Electromagnetic radiation

Introduction

In the previous chapters, we have studied solutions to Maxwell's equations in the form of plane waves propagating in regions that are either empty of conduction charges (vacuum, linear homogeneous and isotropic dielectrics) or may contain free charge and current densities (metals). We discussed propagation, reflection and refraction of the electromagnetic field at an interface between two media, dispersion and absorption. However, we only considered waves propagating away from their emission region and did not investigate how these waves detach from their generation region before propagating.

In this chapter, we are interested in the emission of electromagnetic radiation. In particular, among all the fields generated by a charge or current distribution, we will focus on those capable of detaching from their source and propagating away. We will see that any distribution of charge that varies in time radiates energy in the form of electromagnetic fields. Our first objective will then be to determine the fields when the temporal and spatial dependence of the sources, $\rho(\mathbf{x}, t)$ and $\mathbf{J}(\mathbf{x}, t)$, are known. We limit the discussion to fields generated in vacuum or in linear, homogeneous and isotropic media having permittivity $\epsilon \simeq \epsilon_0$ and permeability $\mu \simeq \mu_0$ close to those in vacuum. This is valid for metals but also a good approximation for air or the atmosphere, and therefore for investigating the propagation in air of electromagnetic waves generated by an antenna.

Electromagnetic radiation is of course the core physical process that allows us to send signals through space between two antennas, one acting as emitter and the other as receiver. While many different types of antennas were developed from the twentieth century, it is essentially due to the fact, among other technical considerations, that an antenna has to be commensurate to the wavelength, but the physical principles governing how antennas work are the same for all types of antennas. In particular emitter and receiver antennas work in the same way. We will provide examples of these general principles.

The same principles will allow us to understand how an incident wave is scattered by small regions of matter with smooth edges, giving rise to a scattered field that is emitted by the charges that are put in motion by the incident field within the matter. The scattered field is superposed to the incident field and the resulting total field has to consistently satisfy Maxwell's equations in matter, including charges and currents. We will see how these concepts can be applied to free electrons and to bound

electrons, which helps us understand natural phenomena such as the color of the sky, and scattering phenomena in general.

12.1 Retarded Potentials - General Solution to Maxwell's Equations

To start with, we will establish the general solution to Maxwell's equation in vacuum, with time dependent current or charge density. This will generalize what we already know from electrostatics and magnetostatics, and will show that electromagnetic radiation exhibits retardation with respect to the source (charge and current), due to the finite propagation velocity of electromagnetic waves.

We have already established that the propagation equations for the fields read

$$\nabla^2 \mathbf{E}(\mathbf{x}, t) - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}(\mathbf{x}, t)}{\partial t^2} = \frac{\nabla \rho(\mathbf{x}, t)}{\epsilon} + \mu_0 \frac{\partial \mathbf{J}(\mathbf{x}, t)}{\partial t} \quad (12.1)$$

for the electric field and

$$\nabla^2 \mathbf{B}(\mathbf{x}, t) - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}(\mathbf{x}, t)}{\partial t^2} = -\mu_0 \nabla \times \mathbf{J}(\mathbf{x}, t), \quad (12.2)$$

for the magnetic field. These equations, derived from Maxwell's equations, involve the six unknown components of the electric and magnetic fields and the source terms $\rho(\mathbf{x}, t)$ and $\mathbf{J}(\mathbf{x}, t)$. However, only four components of the fields are independent since the original Maxwell equations involve 8 scalar equations for 6 field components, meaning that two scalar equations can be viewed as constraints, leaving only 4 independent field-components. To simplify the search for general solutions to Eqs (12.1) and (12.2) that also satisfy the original Maxwell equations, we introduced the scalar and vector potentials, $V(\mathbf{x}, t)$ and $\mathbf{A}(\mathbf{x}, t)$, which indeed amounts to finding 4 scalar quantities. The potentials allow us to determine all field components from

$$\mathbf{B}(\mathbf{x}, t) = \nabla \times \mathbf{A}(\mathbf{x}, t),$$

and

$$\mathbf{E}(\mathbf{x}, t) = -\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} - \nabla V(\mathbf{x}, t).$$

Under the Lorenz gauge, the potentials satisfy:

$$\nabla^2 V(\mathbf{x}, t) - \mu_0 \epsilon_0 \frac{\partial^2 V(\mathbf{x}, t)}{\partial t^2} = -\frac{\rho(\mathbf{x}, t)}{\epsilon_0} \quad (12.3)$$

$$\nabla^2 \mathbf{A}(\mathbf{x}, t) - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{A}(\mathbf{x}, t)}{\partial t^2} = -\mu_0 \mathbf{J}(\mathbf{x}, t) \quad (12.4)$$

Both potentials satisfy an inhomogeneous wave equation, the solution of which guarantees that the electric and magnetic fields are solutions to the original Maxwell equations¹.

¹Solving directly Eqs. (12.1) and (12.2) does not guarantee that $\nabla \cdot \mathbf{B} = 0$ and $\nabla \cdot (\epsilon_0 \mathbf{E}) = \rho$, even if we used these equations to derive the wave equations (12.1) and (12.2)

In the static case for which the sources do not depend on time $\rho = \rho(\mathbf{x})$, $\mathbf{J} = \mathbf{J}(\mathbf{x})$, the well-known Poisson equation is retrieved for V and \mathbf{A}

$$\begin{aligned}\nabla^2 V(\mathbf{x}) &= -\frac{\rho(\mathbf{x})}{\varepsilon_0}, \\ \nabla^2 \mathbf{A}(\mathbf{x}) &= -\mu_0 \mathbf{J}(\mathbf{x}),\end{aligned}$$

which admit the solutions (see chapters 1 and 7)

$$\begin{aligned}V(\mathbf{x}) &= \frac{1}{4\pi\varepsilon_0} \iiint_{\Omega} \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \\ \mathbf{A}(\mathbf{x}) &= \frac{\mu_0}{4\pi} \iiint_{\Omega} \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'.\end{aligned}$$

12.1.1 Retarded potentials - Solution of the inhomogeneous wave equation

Since the three components of the vector potential and the scalar potential satisfy similar equations only involving a different source term, we can find the complete electromagnetic field if we find the formal solution to the scalar inhomogeneous wave equation

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \phi(\mathbf{x}, t) = -4\pi f(\mathbf{x}, t),$$

where $\phi(\mathbf{x}, t)$ represents the scalar potential or any component of the vector potential, and $f(\mathbf{x}, t)$ denotes the corresponding function of space and time that plays the role of the source term. In this aim, we define the Green function that corresponds to the solution when we have a point source at $\mathbf{x} = \mathbf{x}'$ and $t = t'$:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) G(\mathbf{x} - \mathbf{x}', t - t') = -4\pi \delta(t - t') \delta(\mathbf{x} - \mathbf{x}') \quad (12.5)$$

In terms of Green's function, the general solution for $\phi(\mathbf{x}, t)$ takes the form

$$\phi(\mathbf{x}, t) = \int_{-\infty}^{\infty} dt' \iiint_{\mathbb{R}^3} d^3x' G(\mathbf{x} - \mathbf{x}', t - t') f(\mathbf{x}', t')$$

The equation for the Green function can be solved in the Fourier domain. Indeed, if $G(\mathbf{x} - \mathbf{x}', \omega)$ is the Fourier transform with respect to $t - t'$ of $G(\mathbf{x} - \mathbf{x}', t - t')$

$$G(\mathbf{x} - \mathbf{x}', t - t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega G(\mathbf{x} - \mathbf{x}', \omega) e^{-i\omega(t-t')},$$

applying the d'Alembert operator $\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}$ to $G(\mathbf{x} - \mathbf{x}', t - t')$ and commuting it with the integral with respect to ω , we obtain

$$\begin{aligned}\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega G(\mathbf{x} - \mathbf{x}', \omega) e^{-i\omega(t-t')} \right) &= \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) G(\mathbf{x} - \mathbf{x}', \omega) e^{-i\omega(t-t')}.\end{aligned}$$

Using Eq. (12.5), we find

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \left(-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \nabla^2 \right) G(\mathbf{x} - \mathbf{x}', \omega) e^{-i\omega(t-t')} = -4\pi \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega(t-t')} \delta(\mathbf{x} - \mathbf{x}'),$$

that is,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \left\{ \left(\frac{\omega^2}{c^2} + \nabla^2 \right) G(\mathbf{x} - \mathbf{x}', \omega) + 4\pi \delta(\mathbf{x} - \mathbf{x}') \right\} e^{-i\omega(t-t')} = 0,$$

leading to the equation satisfied by the Fourier transform of the Green function

$$(\nabla^2 + k^2) G(\mathbf{x} - \mathbf{x}', \omega) = -4\pi \delta(\mathbf{x} - \mathbf{x}'),$$

where we introduced the (real) variable

$$k \equiv \frac{\omega}{c},$$

which will include the frequency dependence of the Green function. Note the relative similarity of the result with the Laplace equation, where the source is a point charge located at \mathbf{x}' . We know that the solution to this problem has a spherical symmetry with respect to \mathbf{x}' . We will then look for a solution of the form

$$G(\mathbf{x} - \mathbf{x}', \omega) = G(|\mathbf{x} - \mathbf{x}'|, \omega).$$

Calling $r = |\mathbf{x} - \mathbf{x}'|$, we obtain

$$\nabla^2 G(r, \omega) = \frac{1}{r} \frac{d^2}{dr^2} r G(r, \omega),$$

and so

$$\frac{1}{r} \frac{d^2}{dr^2} r G(r, \omega) + k^2 G(r, \omega) = -4\pi \delta(r)$$

whose solution is²

$$G(r, \omega) = \frac{1}{r} e^{\pm ikr}.$$

²The function $rG(r, \omega)$ satisfies a second order ordinary differential equation which is that for a harmonic oscillator with pseudo-time r and pseudo-frequency $k = \omega/c$ for all $r \neq 0$. Hence the solution is a linear superposition of the two components

$$G_{\pm}(r, \omega) = \frac{g_{\pm}}{r} \exp(\pm ikr).$$

The constants g_{\pm} are found from the boundary condition at $r=0$, which is obtained by integrating the original equation

$$(\nabla^2 + k^2) G(r, \omega) = -4\pi \delta(r).$$

over the volume Ω bounded by a sphere of radius r . Since $\nabla^2 G \equiv \nabla \cdot \nabla G$, We find

$$\iiint_{\Omega} \nabla \cdot \nabla G(r, \omega) d^3x = -4\pi \iiint_{\Omega} \delta(r) d^3x - k^2 \iiint_{\Omega} G(r, \omega) d^3x,$$

that is, after using Ostrogradsky's theorem to transform the left-hand side into a surface integral and using the property $\iiint_{\Omega} \delta(r) d^3x = 1$

$$(\nabla G(r, \omega)) \cdot \mathbf{u}_r 4\pi r^2 = -4\pi - k^2 \int_0^r G(r, \omega) 4\pi r^2 dr.$$

For the left-hand side:

$$(\nabla G(r, \omega)) \cdot \mathbf{u}_r 4\pi r^2 = 4\pi r^2 \frac{\partial G}{\partial r} = 4\pi g_{\pm} e^{\pm ikr} (\pm ikr - 1).$$

We then find

$$G(\mathbf{x} - \mathbf{x}', \omega) = \frac{1}{|\mathbf{x} - \mathbf{x}'|} e^{\pm i(\omega/c)|\mathbf{x} - \mathbf{x}'|}$$

An inverse Fourier transform yields

$$\begin{aligned} G(\mathbf{x} - \mathbf{x}', t - t') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\mathbf{x} - \mathbf{x}', \omega) e^{-i\omega(t-t')} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\pm i(\omega/c)|\mathbf{x} - \mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} e^{-i\omega(t-t')} d\omega \\ G(\mathbf{x} - \mathbf{x}', t - t') &= \frac{1}{2\pi |\mathbf{x} - \mathbf{x}'|} \int_{-\infty}^{\infty} d\omega e^{-i\omega(t-t' \mp |\mathbf{x} - \mathbf{x}'|/c)} = \frac{\delta(t - t' \mp |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}. \end{aligned}$$

Green's function is then found in the form of two possible solutions depending on the sign that is chosen:

Green's **retarded function** is obtained by taking the solution with the minus sign

$$G_-(\mathbf{x} - \mathbf{x}', t - t') = \frac{\delta(t - t' - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}$$

The advanced Green function is obtained with the plus sign

$$G_+(\mathbf{x} - \mathbf{x}', t - t') = \frac{\delta(t - t' + |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}.$$

Mathematically, both are solutions to the wave equation. Physically, note that $\frac{|\mathbf{x} - \mathbf{x}'|}{c} = \frac{r}{c}$ represents the time it takes for a signal to propagate at speed c over a distance $r = |\mathbf{x} - \mathbf{x}'|$. Thus the retarded Green function represents the impulse response at point \mathbf{x} and time t to a source located at \mathbf{x}' in the state it had at the retarded time $t' = t - r/c$, that is, the signal at an observation point \mathbf{x} and time t follows the emission by a source a distance r away from \mathbf{x} after a delay corresponding to the signal propagation from the source to the observation point. This solution is in agreement with our intuition that cause precedes effect.

Similarly, the advanced Green function represents the impulse response observed at point \mathbf{x} and time t to a source located at \mathbf{x}' , a distance r away from the observer, in its future state at time $t' = t + r/c$. Although it is a perfectly valid solution from the mathematical point of view, the advanced solution is not causal.

Green's retarded function is causal, and it is the only physically acceptable solution:

$$\begin{aligned} \phi(t, \mathbf{x}) &= \int_{-\infty}^{\infty} dt' \iiint_{\mathbb{R}^3} d^3x' \frac{\delta(t - t' - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} f(t', \mathbf{x}') \\ &= \iiint_{\mathbb{R}^3} d^3x' \frac{f(t - |\mathbf{x} - \mathbf{x}'|/c, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \end{aligned}$$

For the last term on the right-hand side, integrating by parts, we find:

$$-k^2 \int_0^r G(r, \omega) 4\pi r^2 dr = -k^2 4\pi g_{\pm} \int_0^r e^{\pm ikr} r dr = 4\pi g_{\pm} (\pm ikr e^{\pm ikr} - e^{\pm ikr} + 1).$$

Identifying both sides, we finally find

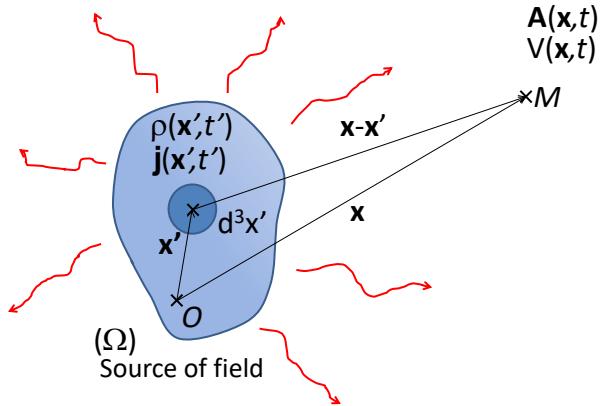
$$g_{\pm} = 1.$$

Finally, going back to the original source terms in the equations governing the scalar and vector potentials, the solutions for the potentials are given by

$$V(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (12.6)$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (12.7)$$

In these expressions, the delay that appears in the time dependency $|\mathbf{x} - \mathbf{x}'|/c$ clearly corresponds to the time it takes light to travel the distance that separates the point of emission \mathbf{x}' from the observation point \mathbf{x} . This solution to Maxwell's equations makes it obvious that electrodynamics is spreading at speed c . This is why, for example, when we observe a star located several light-years from the Earth, this star may have already disappeared. We also see that in Eqs (12.6) and (12.7), the potentials decay as $1/r$, which is necessary to satisfy the conservation of energy, as will be shown below.



12.2 Radiation

The purpose now is to obtain expressions for potentials and fields radiated by an arbitrary charge distribution. We will see that the dipole moment of the distribution plays a fundamental role. We suppose that the charge distribution is confined to a finite volume Ω . While the charge density is time dependent in general, the total charge within Ω is not, in agreement with charge conservation and our assumption that the conduction current at the surface of Ω is zero for the charge to remain localized.

12.2.1 Scalar potential

The scalar potential is given by

$$V(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} d^3x' \frac{\rho(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}.$$

- SCALAR POTENTIAL IN THE NEAR ZONE

Let's see first what happens near the volume Ω , that is, when $|\mathbf{x} - \mathbf{x}'| \ll cT$, where T denotes a time scale for the variation of the charge density. If we consider a sinusoidal signal of frequency ω , then $T = 2\pi/\omega$ and the condition reads $|\mathbf{x} - \mathbf{x}'| \ll \lambda$, defining a near zone as the region at less than a wavelength from the source Ω . Using a Taylor expansion to first order, we can write

$$\rho(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c) \approx \rho(\mathbf{x}', t) - \frac{1}{c} \frac{\partial \rho(\mathbf{x}', t)}{\partial t} |\mathbf{x} - \mathbf{x}'|. \quad (12.8)$$

This expansion leads to

$$V(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \left\{ \iiint_{\Omega} \frac{\rho(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x' - \frac{1}{c} \iiint_{\Omega} \frac{\partial \rho(\mathbf{x}', t)}{\partial t} d^3x' + \dots \right\}$$

The second term is null, given the conservation of the total charge in Ω . Finally, the potential near the source is obtained:

$$V_{\text{near}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} d^3x' \frac{\rho(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|}.$$

We see that in the near zone, the delay that represents the propagation time from the source to the observation point is neglected, and the potential obtained is equal to the static potential.

- SCALAR POTENTIAL IN THE FAR ZONE

Now we consider the opposite case, far enough from Ω , i.e., using an origin close to Ω ,

$$|\mathbf{x}| \gg |\mathbf{x}'|, \quad \forall \mathbf{x}' \in \Omega$$

Writing in spherical coordinates $\mathbf{x} = r\mathbf{u}_r$ leads to the expansion

$$|\mathbf{x} - \mathbf{x}'| = r \sqrt{1 - \frac{2\mathbf{u}_r \cdot \mathbf{x}'}{r} + \frac{|\mathbf{x}'|^2}{r^2}} \approx r \left(1 - \frac{\mathbf{u}_r \cdot \mathbf{x}'}{r} + O\left(\frac{|\mathbf{x}'|^2}{r^2}\right) \right). \quad (12.9)$$

Then, at first order, this yields the following expansion:

$$V_{\text{far}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0 r} \iiint_{\Omega} d^3x' \rho(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c).$$

(12.10)

We see that in the denominator, it is sufficient to retain only the first term r of the expansion, while in the numerator, we retain a correction to the main term. The charge density is evaluated at time $t^* = t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c$, which is simply interpreted as the sum of the retarded time

$$\tau = t - \frac{r}{c}$$

and a correction $\mathbf{u}_r \cdot \mathbf{x}'/c$ that has the order or magnitude of L/c , the time required for a signal to propagate across the source of dimension L at velocity c . This correction is also important as it retains the angular dependence of the fields in the far-zone³.

12.2.2 Vector potential

We can use the same expansions to express the vector potential

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (12.11)$$

in the near zone and in the far zone.

- **VECTOR POTENTIAL IN THE NEAR ZONE**

Without any surprise, an expansion of the current density in Eq. (12.11) similarly to the expansion of the charge density (12.8) in the near zone $|\mathbf{x} - \mathbf{x}'| \ll \lambda$ leads to

$$\mathbf{A}_{\text{near}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (12.12)$$

which is the static vector potential. We see that the expression known from chapter 7 is a particular case of the general solution. Simply allowing the current density to be time-dependent in this expression for the static vector potential leads to a solution in the time-dependent regime that has a validity range limited to the near-zone.

- **VECTOR POTENTIAL IN THE FAR ZONE**

In the far zone $|\mathbf{x}'| \ll |\mathbf{x}| = r$, the expansion (12.9) at leading order yields

$$\mathbf{A}_{\text{far}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\mathbb{R}^3} \mathbf{J}(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c) d^3x'. \quad (12.13)$$

12.2.3 Radiation fields

Radiation fields denote the prevailing terms of the electromagnetic field in the far-zone. As will be seen below, the electromagnetic field in the far-zone decays as $1/r$, in contrast with the Coulomb field in electrostatics for instance. This has important consequences on the transported electromagnetic energy. Before discussing these consequences, expressions for the radiation fields are required. We

³It is clear that without this correction, the scalar potential would have spherical symmetry.

will use the potentials in the far-zone (12.13) and (12.10) to find these expressions by means of the equations linking potentials and fields:

$$\mathbf{E}(\mathbf{x}, t) = -\frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} - \nabla V(\mathbf{x}, t), \quad (12.14)$$

$$\mathbf{B}(\mathbf{x}, t) = \nabla \times \mathbf{A}(\mathbf{x}, t). \quad (12.15)$$

We start with the magnetic field, retaining only the prevailing terms. The final result, demonstrated in Ex. 12.1, reads

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = -\frac{\mu_0}{4\pi c r} \mathbf{u}_r \times \frac{\partial}{\partial t} \left(\iiint_{\Omega} d^3 x' \mathbf{J}(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c) \right). \quad (12.16)$$

Proceeding with the electric radiation field and retaining only the dominant terms varying as $1/r$, the final result, as shown in Ex. 12.2, reads

$$\mathbf{E}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times \left(\mathbf{u}_r \times \frac{\partial}{\partial t} \iiint_{\Omega} d^3 x' \mathbf{J}(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c) \right). \quad (12.17)$$

Note that \mathbf{E}_{rad} and \mathbf{B}_{rad} can be entirely determined from the radiation vector potential

$$\mathbf{A}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \left(\iiint_{\Omega} d^3 x' \mathbf{J}(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c) \right). \quad (12.18)$$

The radiation fields indeed read

$$\begin{aligned} \mathbf{B}_{\text{rad}}(\mathbf{x}, t) &= -\frac{\mathbf{u}_r}{c} \times \frac{\partial}{\partial t} \mathbf{A}_{\text{rad}}(\mathbf{x}, t), \\ \mathbf{E}_{\text{rad}}(\mathbf{x}, t) &= -c \mathbf{u}_r \times \mathbf{B}_{\text{rad}}(\mathbf{x}, t), \end{aligned}$$

or equivalently to the last equation, we can write

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = \frac{\mathbf{u}_r}{c} \times \mathbf{E}_{\text{rad}}(\mathbf{x}, t).$$

The electromagnetic radiation field thus corresponds to a spherical wave due to the dependence of \mathbf{E}_{rad} and \mathbf{B}_{rad} as $1/r$. As will be seen below, this means that the Poynting vector will transport electromagnetic radiation energy undiminished to radial distances arbitrarily far from the source. The facts that $|\mathbf{B}_{\text{rad}}| = |\mathbf{E}_{\text{rad}}|/c$ and $(\mathbf{E}_{\text{rad}}, \mathbf{B}_{\text{rad}}, \mathbf{u}_r)$ forms a right-handed orthogonal triad means that the electromagnetic radiation field has the structure of a quasi-plane wave, propagating along the radial direction.

Example 12.1 - Derivation of the magnetic radiation field from the vector potential (12.18)

Use Eq. (12.18) together with (12.15) to derive the expression Eq. (12.16) for the magnetic radiation field corresponding to the leading order of $\mathbf{B}(\mathbf{x}, t)$ varying as $1/r$ in the far-zone.

Solution

For the magnetic field, we start from the radiation vector potential

$$\mathbf{A}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*)$$

where

$$t^* = t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c}.$$

First use the following vector calculus identity ^a

$$\nabla \times \left(\frac{\iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*)}{r} \right) = \nabla \left(\frac{1}{r} \right) \times \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) + \frac{1}{r} \nabla \times \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*).$$

The first term decays as $(1/r^2)$, while for the second,

$$\nabla \times \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) = \iiint_{\Omega} d^3x' \nabla \times \mathbf{J}(\mathbf{x}', t^*) = \iiint_{\Omega} d^3x' \nabla t^* \times \frac{\partial}{\partial t} \mathbf{J}(\mathbf{x}', t^*).$$

We therefore need to calculate

$$\nabla t^* = -\frac{1}{c} [\nabla r - \nabla(\mathbf{u}_r \cdot \mathbf{x}')] = -\frac{\mathbf{u}_r}{c} + \frac{\mathbf{x}'}{cr} - \frac{\mathbf{x} \cdot \mathbf{x}'}{cr^2} \mathbf{u}_r$$

and we retain only the dominant (first) term on the right-hand side:

$$\boxed{\nabla t^* \approx -\frac{\mathbf{u}_r}{c}}$$

We finally commute the integral and differentiation with respect to time:

$$\nabla \times \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) = -\frac{\mathbf{u}_r}{c} \times \iiint_{\Omega} d^3x' \frac{\partial}{\partial t} \mathbf{J}(\mathbf{x}', t^*) = -\frac{\mathbf{u}_r}{c} \times \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*),$$

which yields the magnetic radiation field Eq. (12.16)

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = -\frac{\mu_0}{4\pi c r} \mathbf{u}_r \times \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*).$$

^aFor any scalar a and vector \mathbf{b} , $\nabla \times a\mathbf{b} = (\nabla a) \times \mathbf{b} + a\nabla \times \mathbf{b}$.

Example 12.2 - Derivation of the electric radiation field from the potentials (12.10) and (12.18)

Use Eqs (12.10) and (12.18) together with (12.14) to find expression (12.17) for the radiative electric field.

Solution

We start with the gradient of the scalar potential:

$$V_{\text{rad}}(\vec{x}, t) = \frac{1}{4\pi\epsilon_0 r} \iiint_{\Omega} d^3x' \rho(\mathbf{x}', t^*),$$

where

$$t^* = t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c}$$

Applying the vector calculus identity for the gradient of a product^a, we find

$$\nabla V_{\text{rad}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \left[\frac{1}{r} \nabla \left(\iiint_{\Omega} d^3x' \rho(\mathbf{x}', t^*) \right) - \left(\frac{\mathbf{u}_r}{r^2} \right) \iiint_{\Omega} d^3x' \rho(\mathbf{x}', t^*) \right].$$

The second term on the right-hand side generates a field that decays as $1/r^2$, and therefore does not contribute to radiation. For the first term, using the chain rule, we have

$$\begin{aligned} \nabla \left(\iiint_{\Omega} d^3x' \rho(\mathbf{x}', t^*) \right) &= \iiint_{\Omega} d^3x' \nabla \rho(\mathbf{x}', t^*) = \iiint_{\Omega} d^3x' \nabla t^* \frac{\partial}{\partial t} \rho(\mathbf{x}', t^*) \\ &\approx \frac{\mathbf{u}_r}{c} \iiint_{\Omega} d^3x' \frac{\partial}{\partial t} \rho(\mathbf{x}', t^*). \end{aligned}$$

Then we use charge conservation to eliminate the charge density in favor of the current density:

$$\frac{\partial}{\partial t} \rho(\mathbf{x}', t^*) = \left[\frac{\partial}{\partial t'} \rho(\mathbf{x}', t') \right]_{t'=t^*} = -[\nabla' \cdot \mathbf{J}(\mathbf{x}', t')]_{t'=t^*}$$

On the other hand, we can use the chain rule to calculate

$$\nabla' \cdot \mathbf{J}(\mathbf{x}', t^*) = [\nabla' \cdot \mathbf{J}(\mathbf{x}', t')]_{t'=t^*} + \nabla' t^* \cdot \frac{\partial}{\partial t'} \mathbf{J}(\mathbf{x}', t^*)$$

and $\nabla' t^* = \mathbf{u}_r/c$. Thus,

$$\frac{\partial}{\partial t} \rho(\mathbf{x}', t^*) = -\nabla' \cdot \mathbf{J}(\mathbf{x}', t^*) + \frac{\mathbf{u}_r}{c} \cdot \frac{\partial}{\partial t} \mathbf{J}(\mathbf{x}', t^*).$$

When we perform the integral over the volume Ω , the term $-\nabla' \cdot \mathbf{J}(\mathbf{x}', t^*)$ involving the divergence is transformed into a surface integral which vanishes because $\mathbf{J}(\mathbf{x}', t^*) = 0$ for $\mathbf{x}' \in \partial\Omega$. Finally,

$$\nabla \left(\iiint_{\Omega} d^3x' \rho(\mathbf{x}', t^*) \right) = -\frac{1}{c^2} \mathbf{u}_r (\mathbf{u}_r \cdot \boldsymbol{\alpha}(\mathbf{x}, t)),$$

where

$$\boldsymbol{\alpha}(\mathbf{x}, t) \equiv \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*), \quad \text{with } t^* = t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c.$$

Another contribution to the electric radiation field comes from

$$\frac{\partial \mathbf{A}_{\text{rad}}}{\partial t} = \frac{\mu_0}{4\pi r} \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) = \frac{\mu_0}{4\pi r} \boldsymbol{\alpha}(\mathbf{x}, t)$$

Collecting the two contributions, we find

$$\nabla V_{\text{rad}} + \frac{\partial \mathbf{A}_{\text{rad}}}{\partial t} = \frac{\mu_0}{4\pi r} \{ -\mathbf{u}_r (\mathbf{u}_r \cdot \boldsymbol{\alpha}(\mathbf{x}, t)) + \boldsymbol{\alpha}(\mathbf{x}, t) \}.$$

We recognize the triple product expansion: $\mathbf{u}_r \times (\mathbf{u}_r \times \boldsymbol{\alpha}) = \mathbf{u}_r (\mathbf{u}_r \cdot \boldsymbol{\alpha}) - \boldsymbol{\alpha}$, hence

$$\mathbf{E}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times (\mathbf{u}_r \times \boldsymbol{\alpha}(\mathbf{x}, t)) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times \left(\mathbf{u}_r \times \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) \right).$$

^aFor two scalar fields a and b , $\nabla(ab) = a\nabla b + b\nabla a$.

12.2.4 Poynting's vector and radiated energy

The electromagnetic power that is radiated far from the source through a spherical surface S of radius r centered on the source is obtained from the flux of the Poynting vector through S . The electromagnetic radiation field has the structure of a quasi-plane wave with \mathbf{E}_{rad} , \mathbf{B}_{rad} , \mathbf{u}_r forming a right-handed orthogonal triad. The Poynting vector is then radial and reads

$$\begin{aligned} \mathbf{\Pi}_{\text{rad}}(\mathbf{x}, t) &= \frac{1}{\mu_0} \mathbf{E}_{\text{rad}}(\mathbf{x}, t) \times \mathbf{B}_{\text{rad}}(\mathbf{x}, t) = \frac{c}{\mu_0} |\mathbf{B}_{\text{rad}}(\mathbf{x}, t)|^2 \mathbf{u}_r \\ &= \frac{1}{\mu_0 c} \left| \mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2 \mathbf{u}_r. \end{aligned}$$

The power radiated by the distribution of charges is obtained by integrating the Poynting vector on a closed surface containing Ω . Considering a sphere of radius r and using the solid angle element $d\Omega$, defined from the surface element $d\mathbf{S}$ as $d\Omega = d\mathbf{S} \cdot \mathbf{u}_r / r^2 = \sin \vartheta d\vartheta d\phi$:

$$P = \iint_S d\mathbf{S}(\mathbf{x}) \cdot \mathbf{\Pi}_{\text{rad}}(\mathbf{x}, t) = \frac{1}{\mu_0 c} \int_0^{4\pi} d\Omega \left| r \mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2$$

Since an expansion in powers of $1/r$ shows that $\mathbf{A}_{\text{rad}}(\mathbf{x}, t)$ varies as $1/r$ at leading order, we have

$$\lim_{r \rightarrow \infty} \left| r \mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2 = f(\vartheta, \phi),$$

that is, the amount of radiated power dP into a differential element of solid angle $d\Omega$, defined as

$$dP = \lim_{r \rightarrow +\infty} \mathbf{u}_r \cdot \mathbf{\Pi}(\mathbf{x}, t) r^2 d\Omega,$$

does not depend on r but only on ϑ and ϕ , the polar and azimuthal angles of spherical coordinates. We can then define a general formula that quantifies the angular distribution of electromagnetic radiation:

$$\frac{dP}{d\Omega} = \frac{1}{c\mu_0} \left| \mathbf{r} \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2. \quad (12.19)$$

Radiative sources and in particular antennas are characterized by their radiation pattern, which informs us about the angular distribution of the radiated electromagnetic power and is entirely determined by Eq. (12.19). An example can be found in the section on antennas 12.4.

Definition 12.1: Electromagnetic radiation

Electromagnetic radiation corresponds to the electromagnetic fields that transport energy arbitrarily far away from their source without attenuation of the total energy observed at distance r . The electromagnetic radiation field that has the structure of a spherical wave with amplitudes of the electric and magnetic fields varying as $1/r$ and becoming quasi-plane waves propagating along \mathbf{u}_r , far from the source, with \mathbf{E}_{rad} , \mathbf{B}_{rad} and \mathbf{u}_r forming a right-handed orthogonal triad. The radiated fields are entirely determined by the radiation vector potential

$$\mathbf{A}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*) \quad (12.20)$$

with

$$t^* = t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c}$$

The magnetic radiation field, reads

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = -\frac{\mathbf{u}_r}{c} \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \quad (12.21)$$

and the electric radiation field reads

$$\mathbf{E}_{\text{rad}}(\mathbf{x}, t) = \mathbf{u}_r \times \left(\mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right) \quad (12.22)$$

and satisfy

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = \frac{\mathbf{u}_r}{c} \times \mathbf{E}_{\text{rad}}(\mathbf{x}, t).$$

Definition 12.2: Total radiation power and its angular distribution

The electromagnetic power radiated through a spherical surface of radius r reads

$$P = \frac{1}{\mu_0 c} \int_0^{4\pi} d\Omega \left| \mathbf{r} \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2 \quad (12.23)$$

where $d\Omega$ denotes the infinitesimal element of solid angle defined from the surface element $d\mathbf{S}$ for the sphere: $d\Omega = d\mathbf{S} \cdot \mathbf{u}_r / r^2 = \sin \vartheta d\vartheta d\phi$, where ϑ and ϕ denote the polar and azimuthal angles of spherical coordinates.

To calculate the angular distribution of radiation power, it is convenient to introduce the *radiation vector*

$$\boldsymbol{\alpha}(\mathbf{x}, t) = \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*)$$

such that

$$\frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} = \frac{\mu_0}{4\pi r} \boldsymbol{\alpha}(\mathbf{x}, t).$$

The angular distribution of radiated power is then given by the formula

$$\frac{dP}{d\Omega} = \frac{\mu_0}{16\pi^2 c} |\mathbf{u}_r \times \boldsymbol{\alpha}(\mathbf{x}, t)|^2.$$

The total power is obtained by integration over the solid angle:

$$P = \frac{\mu_0}{16\pi^2 c} \int_0^{4\pi} d\Omega |\mathbf{u}_r \times \boldsymbol{\alpha}(\mathbf{x}, t)|^2. \quad (12.24)$$

At large distances $r \rightarrow \infty$, the radiation vector $\boldsymbol{\alpha}(\mathbf{x}, t)$ and $\frac{dP}{d\Omega}$ only depend on the spherical angles θ and ϕ . The characterization of the radiation pattern consists in plotting the time average value for $\left\langle \frac{dP}{d\Omega} \right\rangle$ as a function of the polar and azimuthal angles ϑ and ϕ . The pattern exhibits lobes that provide information about the directivity of the antenna.

12.3 Electric dipole radiation

Even after reducing the general solutions to Maxwell's equations expressed in Eqs (12.6) and (12.7) to their far-zone limit, that is, to the potentials for electromagnetic radiation fields (12.10) and (12.13), it is often a difficult task to go further and find analytic expressions of radiation fields for physically relevant current density distributions. Fortunately, simplifications are possible in the case of current distributions whose dimension L remains small compared to the distance cT traveled by the signal during a characteristic evolution time T of the current source. For a harmonic signal, this simply means that the source dimension is smaller than the wavelength of the signal. It is then possible to perform an expansion of the source terms in powers of L/cT . Retaining only the dominant term in such an expansion leads to the laws for electric dipole radiation, valid at distances r from the source satisfying $L \ll \underbrace{cT}_{\equiv \lambda} \ll r$.

To illustrate this procedure, consider the radiation vector

$$\boldsymbol{\alpha}(\mathbf{x}, t) = \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J} \left(\mathbf{x}', t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \right)$$

The current density under the integral can be expanded as

$$\mathbf{J} \left(\mathbf{x}', t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \right) = \mathbf{J}(\mathbf{x}', \tau) + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \frac{\partial}{\partial t} \mathbf{J}(\mathbf{x}', \tau) + \dots$$

where we introduced the retarded time

$$\tau = t - \frac{r}{c}.$$

Note that the order of magnitude for the second term on the right-hand side is L/cT or L/λ for a harmonic current of wavelength λ , justifying the validity range $L \ll \lambda \ll r$ of this expansion. Integrating over the volume Ω , the radiation vector is expressed as

$$\boldsymbol{\alpha}(\mathbf{x}, t) = \frac{d}{dt} \left(\int_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', \tau) \right) + \frac{\mathbf{u}_r}{c} \cdot \frac{d^2}{dt^2} \left(\int_{\Omega} d^3x' \mathbf{x}' \mathbf{J}(\mathbf{x}', \tau) \right) + \dots$$

Similarly, the charge density that appears in the radiation potential (see Eq. (12.10)) can be expanded as

$$\rho \left(\mathbf{x}', t - \underbrace{\frac{r}{c}}_{\tau} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \right) \approx \rho(\mathbf{x}', \tau) + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \frac{\partial}{\partial t} \rho(\mathbf{x}', \tau) + \dots$$

These expansions, after truncation at leading order, will yield the electromagnetic radiation fields.

12.3.1 Scalar potential for the electric dipole radiation

From the expansion of the charge density, we expand the scalar potential as

$$\begin{aligned} V(\mathbf{x}, t) &\approx \frac{1}{4\pi\epsilon_0 r} \left\{ \iiint_{\Omega} \rho(\mathbf{x}', \tau) + \left(\frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \right) \frac{\partial}{\partial t} \rho(\mathbf{x}', \tau) \right\} d^3x' \\ &= \underbrace{\frac{1}{4\pi\epsilon_0 r} \iiint_{\Omega} \rho(\mathbf{x}', \tau) d^3x'}_Q + \frac{1}{4\pi\epsilon_0 cr} \mathbf{u}_r \cdot \frac{d}{dt} \underbrace{\iiint_{\Omega} \mathbf{x}' \rho(\mathbf{x}', \tau) d^3x'}_{\mathbf{p}(t-r/c)} + \dots \end{aligned}$$

where $Q = \iiint_{\Omega} \rho(\mathbf{x}', t) d^3x'$ is the total charge contained in Ω , which is independent of time, and

$$\mathbf{p}(t) \equiv \iiint_{\Omega} \mathbf{x}' \rho(\mathbf{x}', t) d^3x'$$

is the dipole moment of the charge distribution at time t . Finally

$$V(\mathbf{x}, t) \approx \underbrace{\frac{Q}{4\pi\epsilon_0 r}}_{\text{Static}} + \underbrace{\frac{1}{4\pi\epsilon_0 cr} \mathbf{u}_r \cdot \frac{d\mathbf{p}}{dt}(t-r/c)}_{\text{Radiative}}$$

The first term corresponds to a static potential, exactly equal to the electrostatic potential of a point charge Q at the origin. The second term exists as long as the total dipole moment of the distribution depends on time (that is, when the charges move within Ω). Calling $\dot{\mathbf{p}}(t) \equiv d\mathbf{p}/dt$, we write the scalar potential for electric dipole radiation

$$V_{\text{rad,e}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0 cr} \mathbf{u}_r \cdot \dot{\mathbf{p}}(\tau). \quad (12.25)$$

Note again that the $1/r$ -dependence of the radiative potential differs significantly from the $1/r^2$ dependence of the electrostatic potential generated by a dipole charge distribution.

For example, for a dipole at the origin whose magnitude oscillates harmonically in time:

$$\mathbf{p}(t) = \mathbf{p}_0 \cos \omega t,$$

we obtain:

$$\frac{d}{dt} \mathbf{p}(t) = -\mathbf{p}_0 \omega \sin \omega t$$

and so

$$V_{\text{rad}}(\mathbf{x}, t) = -\frac{1}{4\pi\epsilon_0} \mathbf{p}_0 \cdot \mathbf{u}_r \left(\frac{\omega}{rc} \right) \sin[\omega(t - r/c)].$$

12.3.2 Current associated with a dipole

A variation in time of the dipole moment is associated with a current density (Remember the polarization current $\mathbf{J}_P = \frac{\partial \mathbf{P}}{\partial t}$), and therefore, with the generation of a magnetic field. Indeed:

$$\frac{d\mathbf{p}(t)}{dt} = \frac{d}{dt} \iiint_{\Omega} \mathbf{x}' \rho(\mathbf{x}', t) d^3x' = \iiint_{\Omega} \mathbf{x}' \frac{\partial}{\partial t} \rho(\mathbf{x}', t) d^3x'$$

The law of conservation of charge reads $\partial\rho/\partial t + \nabla \cdot \mathbf{J} = 0$, then

$$\frac{d\mathbf{p}(t)}{dt} = - \iiint_{\Omega} \mathbf{x}' (\nabla' \cdot \mathbf{J}(\mathbf{x}', t)) d^3x'.$$

Using the vector calculus identity $\nabla \cdot (x\mathbf{J}) = J_x + x\nabla \cdot \mathbf{J}$, we can rewrite the component along the x -axis as

$$\begin{aligned} \frac{dp_x(t)}{dt} &= - \iiint_{\Omega} x' (\nabla' \cdot \mathbf{J}(\mathbf{x}', t)) d^3x' \\ &= - \iiint_{\Omega} [\nabla' \cdot (x' \mathbf{J}(\mathbf{x}', t)) - J_x(\mathbf{x}', t)] d^3x'. \end{aligned}$$

The first integral over the volume Ω on the right hand side can be transformed into a surface integral by Ostrogradsky's theorem and is identically zero since $\mathbf{J} = 0$ at the frontier $\partial\Omega$. Reproducing this calculation with the y - and z - components of $d\mathbf{p}/dt$, we obtain

$$\frac{d\mathbf{p}(t)}{dt} = \iiint_{\Omega} \mathbf{J}(\mathbf{x}', t) d^3x'.$$

12.3.3 Vector potential for electric dipole radiation

Now we calculate the vector potential associated with the current density in Ω :

$$\mathbf{A}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\Omega} \mathbf{J} \left(\mathbf{x}', t - \underbrace{\frac{r}{c}}_{\tau} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c} \right) d^3x' \approx \frac{\mu_0}{4\pi r} \iiint_{\Omega} \mathbf{J}(\mathbf{x}', \tau) d^3x'$$

We see that we can use the relation

$$\iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', \tau) = \frac{d}{dt} \mathbf{p}(\tau) \equiv \dot{\mathbf{p}}(\tau),$$

so that the radiation vector potential for the electric dipole radiation reads

$$\mathbf{A}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \dot{\mathbf{p}}(\tau) \quad (12.26)$$

where $\tau = t - r/c$ and, noting that $\mathbf{p}(t)$ is a function of a single variable, $\dot{\mathbf{p}} \equiv \frac{d\mathbf{p}}{dt}$. Hence, the potentials (12.25) and (12.26) involve $\dot{\mathbf{p}}$ evaluated at the retarded time $t - r/c$.

12.3.4 Fields for electric dipole radiation

The electric dipole radiation corresponds to the leading order term in the expansion for the radiation vector $\alpha(\mathbf{x}, t)$ in powers of L/λ . In this case, using the link between the current density and the dipole moment, we have

$$\alpha(\mathbf{x}, t) = \frac{d}{dt} \frac{d\mathbf{p}(\tau)}{dt} + \dots \approx \frac{d^2\mathbf{p}(\tau)}{dt^2} = \ddot{\mathbf{p}}(\tau).$$

We see that this truncation makes the radiation vector dependent only on time. We can then use Eq. (12.22) to find the electric radiation field:

$$\mathbf{E}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times (\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c)) \quad (12.27)$$

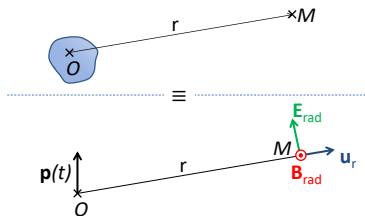
and Eq. (12.21) will in turn lead to the magnetic radiation field

$$\mathbf{B}_{\text{rad,e}}(\mathbf{x}, t) = -\frac{\mu_0}{4\pi c r} (\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c)) \quad (12.28)$$

Note that $\mathbf{E}_{\text{rad,e}}$ and $\mathbf{B}_{\text{rad,e}}$ still correspond to spherical waves^a which satisfy

$$\mathbf{B}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mathbf{u}_r}{c} \times \mathbf{E}_{\text{rad,e}}(\mathbf{x}, t).$$

The radiation electromagnetic field far from the source, when the phase front flattens, keeps the structure of a quasi-plane wave, propagating along the radial direction.



^aThe propagation direction is radial, the wavefronts coincide with the surface of a sphere and the amplitude varies as $1/r$.

12.3.5 Poynting vector and radiation power for electric dipole radiation

Using the vector radiation $\alpha(\mathbf{x}, t) = \ddot{\mathbf{p}}(\tau)$, the angular distribution of radiated power in the case of electric dipole radiation is easy to evaluate from the general formula (12.19):

$$\left[\frac{dP}{d\Omega} \right]_e = \frac{\mu_0}{16\pi^2 c} |\mathbf{u}_r \times \alpha(\mathbf{x}, t)|^2 = \frac{\mu_0}{16\pi^2 c} |\mathbf{u}_r \times \ddot{\mathbf{p}}(\tau)|^2.$$

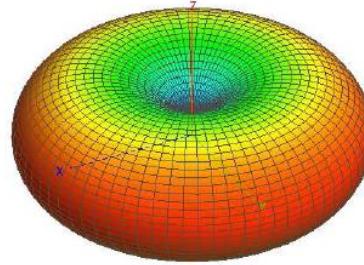
If we chose the axes of the spherical basis of coordinates such that the dipole moment is orientated along the z -axis, we have

$$|\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c)|^2 = |\ddot{\mathbf{p}}(t - r/c)|^2 \sin^2 \vartheta$$

and

$$\left[\frac{dP}{d\Omega} \right]_e = \frac{\mu_0}{16\pi^2 c} |\ddot{\mathbf{p}}(t - r/c)|^2 \sin^2 \vartheta.$$

The radiation pattern displayed in the figure below exhibits an azimuthally symmetric lobe orthogonal to the dipole, that is, characterizing a vertical dipole antenna emitting in all directions but with maximum power directed in a horizontal plane perpendicular to the antenna.



The total radiated power for an electric dipole can be evaluated by integration over the azimuthal and polar angles:

$$P = \frac{\mu_0}{(4\pi)^2 c} |\ddot{\mathbf{p}}(t - r/c)|^2 \underbrace{\int_0^{2\pi} d\varphi}_{2\pi} \underbrace{\int_0^\pi d\vartheta \sin^3 \vartheta}_{4/3}$$

Finally

$$P = \frac{\mu_0}{6\pi c} |\ddot{\mathbf{p}}(t - r/c)|^2 = \frac{1}{6\pi\epsilon_0 c^3} |\ddot{\mathbf{p}}(t - r/c)|^2 \quad (12.29)$$

At any distance r far from the source, the electromagnetic power transported through a spherical surface of radius r only depends on time. This power is transported undiminished at arbitrarily large distances, thanks to the $1/r$ dependence of the radiation fields.

12.3.6 Quasi-plane wave approximation

It is instructive to note that without knowing the general expressions (12.17) and (12.16), the radiative fields (12.27) and (12.28) can be derived in a simple way from the potentials (12.25) and (12.26), using

the *quasi-plane wave approximation*, which consists in considering the amplitude of the radiated electric and magnetic fields as quasi uniform, that is, in keeping r constant in the denominator of (12.25) and (12.26) when taking the curl.

Indeed, starting with Eq. (12.26) and taking $\nabla \times \mathbf{A}_{\text{rad},e}(\mathbf{x}, t)$, we find

$$\mathbf{B}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \nabla \times \dot{\mathbf{p}}(\tau) + \frac{\mu_0}{4\pi} \underbrace{\left(\nabla \frac{1}{r} \right)}_{-1/r^2} \times \dot{\mathbf{p}}(\tau),$$

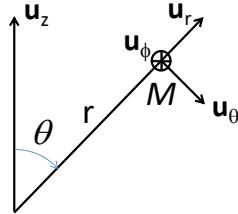
where the first term prevails over the second. This becomes clear if we remember that $\nabla \times \dot{\mathbf{p}}(\tau)$ involves a derivative with respect to r of a function that depends on $\tau = t - r/c$. Its order of magnitude is then

$$\left| \frac{\partial}{\partial r} \dot{\mathbf{p}}(\tau) \right| = \left| \underbrace{\frac{\partial \tau}{\partial r}}_{-1/c} \underbrace{\ddot{\mathbf{p}}(\tau)}_{\sim \dot{\mathbf{p}}/T} \right| = \frac{1}{cT} |\dot{\mathbf{p}}(\tau)| \sim \frac{|\dot{\mathbf{p}}(\tau)|}{\lambda},$$

where $\dot{\mathbf{p}}/T$ is the order of magnitude for $\ddot{\mathbf{p}}$. If we assume a signal dominated by a sinusoidal signal of frequency $\omega = 2\pi/T$, whose wavelength is $\lambda = cT$, the order of magnitude is $\dot{\mathbf{p}}/\lambda$. Hence the ratio of the second to the first term in our expression for the magnetic field scales as λ/r , showing that the first term prevails for $r \gg \lambda$, that is, in the region farther from the source than the wavelength of the signal. The magnetic radiated field then writes

$$\mathbf{B}_{\text{rad},e}(\mathbf{x}, t) \approx \frac{\mu_0}{4\pi r} \nabla \times \dot{\mathbf{p}}(\tau).$$

For the calculation of $\nabla \times \dot{\mathbf{p}}(\tau)$, we can use spherical coordinates (r, θ, ϕ) . A rigorous derivation would require the expression for the curl in spherical coordinates. However, this expression involves terms of higher order varying as $1/r$, which can be neglected with respect to the dominant terms. The result is then the same as that obtained by assuming that $\mathbf{u}_r, \mathbf{u}_\theta, \mathbf{u}_\phi$ is a quasi-Cartesian basis.



The radiated magnetic field then reads

$$\mathbf{B}_{\text{rad},e}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \begin{vmatrix} \partial_r & p'_r(\tau) \\ 0 & p'_\theta(\tau) \\ 0 & p'_\phi(\tau) \end{vmatrix} = \frac{\mu_0}{4\pi r} \mathbf{u}_r \partial_r \times \dot{\mathbf{p}}(\tau)$$

Finally differentiating with respect to r using $\partial_r = -(1/c)\partial_\tau$, we find

$$\mathbf{B}_{\text{rad},e}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \left(-\frac{\mathbf{u}_r}{c} \times \ddot{\mathbf{p}}(\tau) \right)$$

The radiated electric field can then be calculated from Ampère-Maxwell's equation

$$\nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t},$$

by performing consistently the same approximation to evaluate the curl

$$\begin{aligned} \nabla \times \mathbf{B}_{\text{rad,e}} &= \frac{\mu_0}{4\pi r} \underbrace{(\mathbf{u}_r \partial_r)}_{-\frac{\mathbf{u}_r}{c} \frac{\partial}{\partial \tau}} \times \left(-\frac{\mathbf{u}_r}{c} \times \ddot{\mathbf{p}}(\tau) \right) = \frac{1}{c^2} \frac{\partial \mathbf{E}_{\text{rad,e}}}{\partial t} \end{aligned}$$

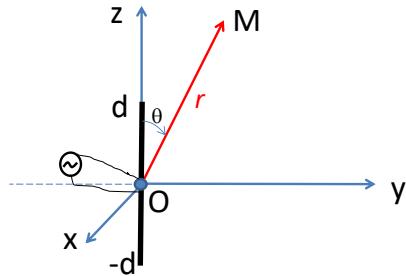
Finally integrating in time leads to

$$\mathbf{E}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times (\mathbf{u}_r \times \ddot{\mathbf{p}}(\tau)) = -\mathbf{u}_r \times c \mathbf{B}_{\text{rad,e}}(\mathbf{x}, t).$$

We retrieve the structure of a plane wave, in agreement with our assumption at the beginning of this section. Summarizing, the shortest way to derive the electromagnetic radiation fields for electric dipole radiation is to start from the expression for the vector potential (12.26), take the curl using the quasi-plane wave approximation to find the magnetic field, and finally use Ampère-Maxwell's equation to find the electric field from the magnetic field.

12.4 Thin-wire antennas

Thin-wire antennas constitute prototypical radiating systems. Incident waves on a thin wire conductor induce a time dependent current in the wire, which in turn drives a voltage across a load resistance. The same principles govern emission of electromagnetic waves, that is, a voltage applied to the terminals of a thin wire conductor drives a time dependent current that will radiate an electromagnetic wave.



Consider a straight, thin-wire antenna of length $2d$ driven at its center by a sinusoidal input voltage of frequency ω . We will consider complex notations and write the current induced in the conductor by this voltage as $\underline{I}(z, t) = \text{Re}[\underline{I}(z, t)]$ and

$$\underline{I}(z, t) = I_0 \sin k(d - |z|) e^{-i\omega t} \quad \text{for } -d \leq z \leq d,$$

where $k = \omega/c$. We assume that the radius of the wire is infinitesimally small so that we can calculate the radiation vector potential from the retarded potential formula

$$\underline{\mathbf{A}}_{\text{rad}} = \frac{\mu_0}{4\pi} \iiint_{\Omega} d^3x' \frac{\mathbf{J}(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} \approx \frac{\mu_0 \mathbf{u}_z}{4\pi r} \int_{-d}^d dz' \underline{I}(z', t - r/c + \mathbf{u}_r \cdot (z' \mathbf{u}_z)/c).$$

Since we wish to characterize the antenna with its radiation pattern, we retained the signal travel time across the source, that is $\mathbf{u}_r \cdot \mathbf{x}'/c$, as a correction to the retarded time. Now using a basis of spherical coordinates, the current at retarded time reads

$$\underline{I}(z', t - r/c + \mathbf{u}_r \cdot (z' \mathbf{u}_z)/c) = I_0 \sin k(d - |z'|) e^{ikrz' \cos \theta} e^{ikr} e^{-i\omega t}$$

and the radiated vector potential,

$$\underline{\mathbf{A}}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0 I_0 \mathbf{u}_z}{4\pi} \frac{e^{ikr - i\omega t}}{r} 2 \int_0^d \cos(k \cos \theta z') \sin[k(d - z')] dz'.$$

The integral can be calculated using first a trigonometric formula⁴. We find

$$\underline{\mathbf{A}}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0 I_0}{2\pi} \frac{e^{ikr - i\omega t}}{kr} \frac{[\cos(kd \cos \theta) - \cos(kd)]}{\sin^2 \theta} \mathbf{u}_z.$$

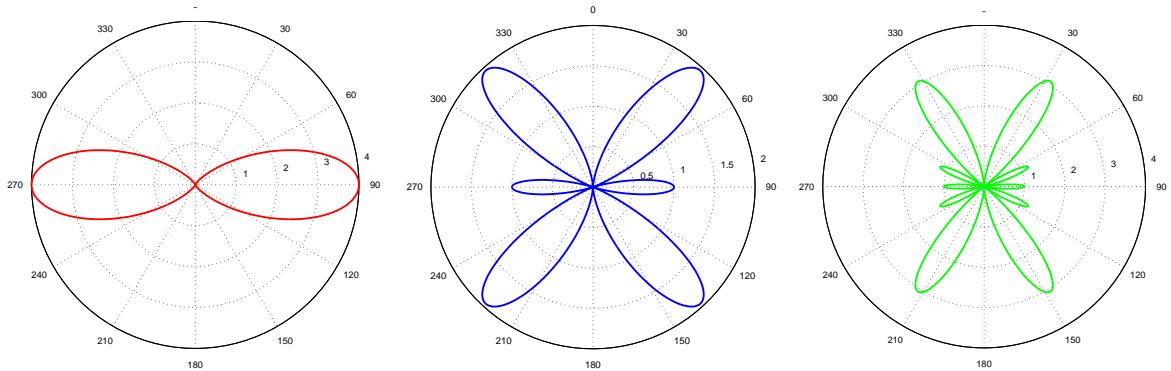
We aim at finding the radiation pattern of the antenna, which informs us about the angular distribution of the radiated electromagnetic power. We know that the angular distribution of radiated power is given by time-average of the quantity

$$\frac{dP}{d\Omega} = \frac{1}{c\mu_0} \left| r \mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2.$$

Introducing the expression of $\mathbf{A}_{\text{rad}}(\mathbf{x}, t)$ for the thin-wire antenna, we calculate the angular distribution of radiated power:

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{1}{2c\mu_0} |\mathbf{r}|^2 | -i\omega \mathbf{A}_{\text{rad}}|^2 \sin^2 \theta = \frac{1}{2} \frac{\mu_0 c I_0^2}{(2\pi)^2} \frac{[\cos(kd \cos \theta) - \cos(kd)]^2}{\sin^2 \theta},$$

where the multiplicative factor 1/2 takes care of the time average procedure for the harmonic signal⁵.



⁴ $\sin \alpha \cos \beta = (1/2)[\sin(\alpha + \beta) + \sin(\alpha - \beta)]$.

⁵ $\langle \cos^2(\omega t - kr) \rangle = 1/2$

The figures show the radiation pattern obtained for different lengths of the antenna: $2d = (2n - 1)\lambda/2$ with $n = 1, 2, 3$. The vertical axis is the z -axis, coinciding with the axis of the antenna. There is revolution symmetry around the z axis (no dependence upon ϕ). The patterns show that there is a single lobe pointing in all horizontal directions when the length of the antenna is equal to $\lambda/2$. Increasing the length of the antenna at fixed wavelength⁶ shows that the number of lobes is equal to the number of half-wavelengths fitting the antenna. and that the main lobe becomes narrower and closer to the axis of the antenna.

12.5 Scattering

An incident electromagnetic wave on a medium induces a motion of charges. These charges are secondary sources of radiation, thus generating a scattered wave, which can have the same frequency as that of the incident wave (elastic diffusion), but it can also have a different frequency (diffusion is then said to be inelastic).

12.5.1 Thomson scattering

Thomson scattering denotes the phenomenon of scattering occurring to an electromagnetic wave interacting with a free electron.

Consider a free electron of charge $q = -e$ of mass m that undergoes the action of a monochromatic electromagnetic wave of the form

$$\mathbf{E}(\mathbf{r}, t) = \operatorname{Re} \left\{ \underline{\mathbf{E}} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \right\}.$$

Under the action of this field, the electron will oscillate at a frequency $f = \omega/2\pi$, thus behaving like an oscillating dipole that radiates at the same frequency as that of the incident field, thus creating an elastic diffusion field. If \mathbf{r} represents the position of the electron, we have:

$$m_e \frac{d^2 \mathbf{r}}{dt^2} = q[\mathbf{E}(\mathbf{r}, t) + \mathbf{v} \times \mathbf{B}(\mathbf{r}, t)]$$

If the speed of the electron is much less than the speed of light c , we can disregard the magnetic force since, for an electromagnetic wave, $|\mathbf{B}| = |\mathbf{E}|/c$. In addition, if the wavelength of the incident field is much greater than the oscillation amplitude of the electron, we can neglect the spatial variation of $\mathbf{E}(\mathbf{r}, t)$ and write $\mathbf{E}(\mathbf{r}, t) = \operatorname{Re} \{ \underline{\mathbf{E}} e^{-i\omega t} \}$. Looking for a stationary solution of the form:

$$\mathbf{r}(t) = \operatorname{Re} \left\{ \underline{\mathbf{r}}_0 e^{-i\omega t} \right\},$$

we have:

$$\operatorname{Re} \left\{ -m_e \omega^2 \underline{\mathbf{r}}_0 e^{-i\omega t} \right\} = \operatorname{Re} \left\{ q \underline{\mathbf{E}} e^{-i\omega t} \right\}.$$

Then, the amplitude of the electron motion is given by:

$$\underline{\mathbf{r}}_0 = -\frac{q}{m_e \omega^2} \underline{\mathbf{E}} \quad (12.30)$$

⁶This is perfectly fine as we did not use the expansion in powers of L/λ in this section.

• SCATTERING CROSS SECTION

The acceleration of an electron that oscillates under the influence of an electromagnetic wave at frequency ω is then given by:

$$\mathbf{a} = \frac{q}{m_e} \mathbf{E} \cos(\omega t)$$

and the power radiated throughout space is:

$$P = \frac{\mu_0 q^2}{6\pi c} \frac{q^2 |E|^2}{m_e^2} \cos^2(\omega t),$$

the time average of which is:

$$P = \frac{\mu_0 q^2}{6\pi c} \frac{|E|^2}{2m_e^2},$$

which can be rewritten as

$$P = \frac{q^2}{4\pi\epsilon_0} \frac{|E|^2}{3m_e^2 c^3} = \frac{4\pi}{3} r_c^2 \epsilon_0 c |E|^2, \quad (12.31)$$

where $r_c = q^2 / (4\pi\epsilon_0 m_e c^2)$ is the classic electron radius. In expression (12.31), we recognize the time average of the Poynting vector of the incident plane wave $\epsilon_0 c |E|^2 / 2$. The diffusion power can therefore be written as the product of an energy flux per unit area (modulus of the incident Poynting vector) multiplied by a quantity that is homogeneous to a surface, called **cross section**.

Definition 12.3: Scattering cross section

The scattering cross section σ corresponds to the surface such that, exposed to the incident flux of an electromagnetic wave, it intercepts a flux of energy equal to the flux of the radiation field, which is the long distance part of the scattered electromagnetic wave. The effective cross section σ_e of a free electron, called the Thomson effective cross section, is obtained from Eq. (12.31) and is given by:

$$\sigma_e = \frac{8\pi}{3} r_c^2$$

An electron is thus seen by an electromagnetic wave as a circular surface of radius similar to the classical radius $r_c \sim 10^{-15}$ m.

12.5.2 Scattering from an atom or a molecule

Atoms are made up of positive charges (protons), whose mass is three orders of magnitude greater than that of the electron. This means that the square of the acceleration, and then the radiated power, is 10^6 times less than the power radiated by the electrons. It is then reasonable to consider only the energy radiated by the electrons of an atom when it interacts with an electromagnetic wave. We can take into account the fact that the electrons are bound to the atomic nucleus by an elastic force. We therefore anticipate a strong dependence of scattering upon the frequency of the incident wave, due to the presence of a resonance frequency ω_0 corresponding to the natural frequency of oscillation of the electron around the nucleus. We will use the Lorentz model that treats the electron-atom system

as a damped harmonic oscillator. Neglecting the magnetic force and assuming an elastic force and a friction force acting on the electron around the nucleus, Newton's equation for the electron reads

$$m \frac{d^2 \mathbf{r}}{dt^2} = q \mathbf{E} - m\gamma \frac{d\mathbf{r}}{dt} - m\omega_0^2 \mathbf{r}$$

where \mathbf{r} is the displacement of the electron with respect to the equilibrium position. Again, assuming a monochromatic incident wave with a much longer wavelength than the oscillation amplitude of the electrons, we look for a solution of the form $\mathbf{r} = \text{Re} \{ \mathbf{r}_0 e^{-i\omega t} \}$ and we have:

$$\text{Re} \left\{ -m\omega^2 \underline{\mathbf{r}}_0 e^{-i\omega t} \right\} = \text{Re} \left\{ q \underline{\mathbf{E}} e^{-i\omega t} + i\omega m\gamma \underline{\mathbf{r}}_0 e^{-i\omega t} - m\omega^2 \underline{\mathbf{r}}_0 e^{-i\omega t} \right\}.$$

Then

$$\underline{\mathbf{r}}_0 = -\frac{q}{m} \frac{\underline{\mathbf{E}}}{\omega^2 - \omega_0^2 + i\gamma\omega}. \quad (12.32)$$

• SCATTERING CROSS SECTION

Since the radiation power is proportional to the square of the acceleration, and as in a harmonic motion, the latter is proportional to the motion amplitude \mathbf{r}_0 , we have:

$$\frac{\sigma}{\sigma_e} = \frac{|\underline{\mathbf{r}}_0|^2}{|\underline{\mathbf{r}}_0^{\text{free}}|^2},$$

where $\underline{\mathbf{r}}_0^{\text{free}}$ corresponds to the result found previously for a free electron - See Eq. (12.30). We have:

$$\frac{\sigma}{\sigma_e} = \frac{\omega^4}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2}$$

so that the cross section of an electron that is elastically bound to an atom reads

$$\sigma(\omega) = \sigma_e \frac{\omega^4}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} \quad (12.33)$$

• THOMSON SCATTERING $\omega \gg \omega_0, \gamma$

In the case of high frequencies such that $\omega \gg \omega_0, \gamma$, we retrieve Thomson scattering since $\sigma(\omega) \approx \sigma_e$. When the electromagnetic wave has a frequency much larger than vibration the frequencies ω_0 of electrons in atoms, electrons then behave as free electrons. Diffraction by X-rays is typically well described by Thomson scattering.

• RESONANT SCATTERING

We see from Eq. (12.33) that the interaction between the electromagnetic wave and an electron bound to an atom is particularly strong when the frequency of the wave is in resonance with the vibration frequency of the electron in the atom, ω_0 . From Eq. (12.33), we see that in the limit $\omega \approx \omega_0$

$$\begin{aligned}\sigma(\omega) &= \sigma_e \frac{\omega^4}{[(\omega - \omega_0)(\omega + \omega_0)]^2 + \gamma^2 \omega^2} \\ &\approx \sigma_e \frac{\omega_0^4}{[(\omega - \omega_0)(2\omega_0)]^2 + \gamma^2 \omega_0^2} \\ \sigma(\omega) &\approx \sigma_e \frac{\omega_0^2}{4(\omega - \omega_0)^2 + \gamma^2}\end{aligned}$$

The dependence of $\sigma(\omega)$ on ω around ω_0 is approximately a Lorentzian of width $1/\gamma$. The maximum value of $\sigma(\omega)$ is obtained for $\omega = \omega_0$:

$$\sigma(\omega_0) = \sigma_e \left(\frac{\omega_0}{\gamma} \right)^2.$$

In the visible domain, $\omega_0 \sim 10^{15} \text{ s}^{-1}$, and $\gamma = 2/3\omega_0\tau \sim 10^8 \text{ s}^{-1}$, so that the cross section can be many orders of magnitude larger than that of the free electron, $\sigma(\omega_0) \sim 10^{14} \times \sigma_e$.

• RAYLEIGH SCATTERING $\omega \ll \gamma \ll \omega_0$

If $\omega \ll \gamma \ll \omega_0$, we are in the case of Rayleigh scattering. The scattering cross section in this case becomes

$$\sigma(\omega) = \sigma_e \frac{\omega^4}{\omega_0^4}.$$

The scaling as ω^4 shows that scattering becomes inefficient at low frequencies toward the far-infrared domain.

Rayleigh scattering is responsible for the blue color of the sky. The ω^4 scaling of the cross section means that blue color in the visible spectrum, with frequency twice larger than the red color, is scattered 16 times more than red when the white light from the Sun encounters atoms in the atmosphere.

Rayleigh scattering is also responsible for the red color of the sky at sunset, when the thickness of the atmosphere crossed by the light is maximum leading to a much more efficient scattering of blue light compared to red light

12.6 Exercises

Example 12.3 - Generation of Plane Waves

We have seen important solutions to Maxwell's equations in the form of plane electromagnetic waves. They are very useful because they are mathematically simple, and many electrodynamic phenomena can be described from this type of waves. We mentioned that plane waves are only an approximation (for example, they can represent spherical waves far from their sources). One could ask if it is possible to generate real plane waves from a distribution of charge and current? We will see that monochromatic plane waves are generated by an infinite planar distribution of charge that oscillates with frequency ω . Clearly this distribution cannot be generated in practice but a sufficiently large planar surface could be a source of waves that approximate the

ideal plane waves obtained from an infinite surface.

Consider a charge distribution in the xy -plane. The charge density can be written as

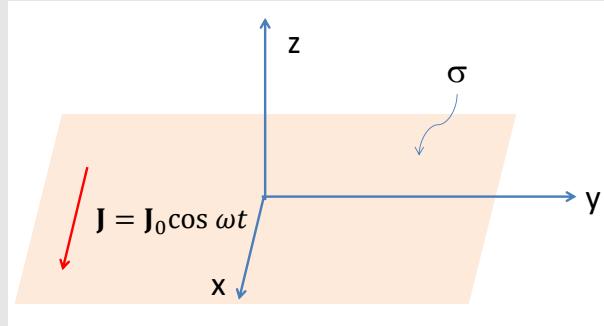
$$\rho(t, \mathbf{x}) = \sigma \delta(z)$$

with σ a surface charge density that is a static; it does not depend on time. Now, suppose the charges oscillate in the x -direction, so that the velocity of one of them can be written as

$$\mathbf{v}(t, \mathbf{x}) = v_0 \cos \omega t \mathbf{u}_x$$

This velocity does not depend on the position of the charge in the plane (they all oscillate in the same way). Hence the associated current density reads

$$\mathbf{J}(t, \mathbf{x}) = \sigma v_0 \delta(z) \cos \omega t \mathbf{u}_x$$



Note that this is consistent with the law of conservation of charge

$$\nabla \cdot \mathbf{J}(\mathbf{x}, t) = 0 = -\frac{\partial \rho(\mathbf{x}, t)}{\partial t}.$$

The vector potential corresponding to this current distribution is given by

$$\mathbf{A}(t, \mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} d^3 x' \int dt' \mathbf{J}(t', \mathbf{x}') \frac{\delta(t - t' - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|},$$

that is,

$$\mathbf{A}(t, \mathbf{x}) = \frac{\mu_0}{4\pi} \sigma \mathbf{u}_x \int dt' \iiint_{\mathbb{R}^3} d^3 x' \delta(z') \frac{\delta(t - t' - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}.$$

Now, define

$$\Gamma(t - t', \mathbf{x}) = \iiint_{\mathbb{R}^3} d^3 x' \delta(z') \frac{\delta(t - t' - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|}.$$

Integrating over z' , we find

$$\Gamma(t - t', \mathbf{x}) = \iint_{\mathbb{R}^2} dx' dy' \frac{\delta(t - t' - \frac{1}{c} \sqrt{(x - x')^2 + (y - y')^2 + z^2})}{\sqrt{(x - x')^2 + (y - y')^2 + z^2}}$$

It is clear, given the infinite charge distribution on the plane and its invariance by translation along the x - or y -axis, that this integral will only depend on z . For this reason, it is convenient to evaluate this expression on the axis z , that is to say at $\mathbf{x} = (0, 0, z)$

$$\Gamma(t - t', \mathbf{x}) = \iint_{\mathbb{R}^2} dx' dy' \frac{\delta\left(t - t' - \frac{1}{c}\sqrt{x'^2 + y'^2 + z^2}\right)}{\sqrt{x'^2 + y'^2 + z^2}}$$

Using the change of variable from Cartesian x', y' to polar l, θ coordinates

$$\begin{cases} x' = l \cos \theta \\ y' = l \sin \theta \end{cases} \quad \text{and} \quad dx' dy' = ldld\theta.$$

we find

$$\Gamma(t - t', z\mathbf{u}_z) = 2\pi \int_0^\infty \frac{\delta(t - t' - \sqrt{l^2 + z^2}/c)}{\sqrt{z^2 + l^2}} l dl = 2\pi \int_0^\infty \frac{\delta(t - t' - r/c)}{r} l dl,$$

where $r \equiv \sqrt{z^2 + l^2}$. Now we make a second change of variable

$$s = \frac{r}{c} - (t - t'), \quad \rightarrow \quad ds = \frac{ldl}{cr} = \frac{ldl}{c^2(s + (t - t'))}.$$

Hence

$$\Gamma(t - t', z\mathbf{u}_z) = 2\pi c \int_{|z|/c - (t - t')}^\infty ds \delta(s) = 2\pi c \Theta(t - t' - |z|/c),$$

where $\Theta(t)$ denotes the Heavyside function^a, so that

$$\begin{aligned} \mathbf{A}(t, z) &= \frac{\mu_0}{4\pi} \sigma v_0 2\pi c \mathbf{u}_x \int_{-\infty}^\infty dt' \Theta(t - t' - |z|/c) \cos \omega t' \\ &= \frac{\mu_0}{2} \sigma v_0 c \mathbf{u}_x \int_{-\infty}^{t - |z|/c} dt' \cos \omega t' = \frac{\mu_0}{2\omega} \sigma v_0 c \mathbf{u}_x \sin[\omega(t - |z|/c)] \end{aligned}$$

Finally

$$\mathbf{A}(t, z) = \frac{E_0}{\omega} \sin [\omega(t - |z|/c)] \mathbf{u}_x$$

with $E_0 = \frac{1}{2}\mu_0\sigma v_0 c$. The solution is in the form

$$\mathbf{A}(t, z) = \frac{E_0}{\omega} \sin \omega(t - \operatorname{sgn}(z)z/c) \mathbf{u}_x,$$

where

$$\operatorname{sgn}(z) = \begin{cases} 1 & \text{if } z > 0, \\ -1 & \text{if } z < 0. \end{cases}$$

It is a plane wave that travels with velocity c in the positive z -direction for $z > 0$ and negative z -direction for $z < 0$.

The electric field of radiation is given by

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = -E_0 \cos \omega(t - |z|/c) \mathbf{u}_x$$

and the magnetic field

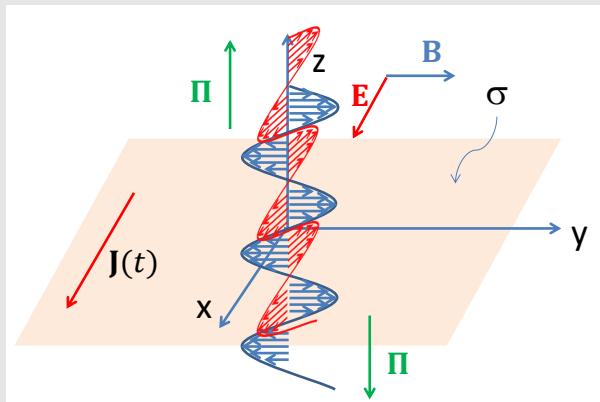
$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{E_0}{c} \epsilon(z) \cos \omega(t - |z|/c) \mathbf{u}_y,$$

where $\text{sgn}(z) = 1$ for $z > 0$ and $\epsilon(z) = -1$ for $z < 0$. Note that

$$\frac{|B|}{|E|} = \frac{1}{c}$$

Furthermore, \mathbf{E} is perpendicular to \mathbf{B} , and both fields are perpendicular to the direction of propagation z (transverse waves). The Poynting vector is given by

$$\Pi(t, \mathbf{x}) = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} = \text{sgn}(z) \epsilon_0 c E_0^2 \cos^2[\omega(t - |z|/c)] \mathbf{u}_z$$



Energy that propagates in the z direction.

Note

Radiative fields of a single point charge are proportional to the acceleration of the charge. Here, we have obtained a very different result as the radiation fields of the oscillating plane distribution of charge are proportional to the velocities of the charges. However, there is no contradiction, since when the radiation fields due to each individual charge are added up, the same result is recovered (see Feynman's lectures on Physics, Vol 1, Ch 30)

^aThe Heavyside function is actually not a function but a distribution, which can be seen as a step function $\Theta(t) = 0$ for $t < 0$, $\Theta(0) = 1/2$, $\Theta(t) = 1$ for $t > 0$, and yields the Dirac distribution by differentiation in the sense of distributions: $\Theta'(t) = \delta(t)$.

Example 12.4 - Radiation of the hydrogen atom

Consider an electron that describes a circular orbit of radius $r = a_0$ around the proton, where $a_0 = 0.529 \text{ \AA}$ is the Bohr radius. Since the electron accelerates, it emits an electromagnetic radiation during its circular motion.

- a. What is the total energy ΔE radiated by the electron during the period of a revolution around the nucleus?
- b. What is the ratio between ΔE and the mechanical energy E of the electron in the circular orbit?
- c. If the electron is initially at a distance of $r_B = 0.529 \text{ \AA}$ (the Bohr radius) from the proton, how much time is required for the electron to collapse at the center of the nucleus?

Solution

- a. The balance of forces for the electron yields:

$$\mathbf{F}(t) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{r^2} \mathbf{u}_r = m_e \mathbf{a}(t)$$

Then the acceleration of the electron is given by:

$$\mathbf{a}(t) = -\frac{e^2}{4\pi\epsilon_0 m_e r^2} \mathbf{u}_r$$

Thus, the power emitted by the electron is given by

$$P = \left(\frac{\mu_0 e^2}{6\pi c} \right) |\mathbf{a}|^2 = \left(\frac{\mu_0}{6\pi c} \right) \frac{e^6}{(4\pi\epsilon_0)^2 m_e^2 r^4}$$

The energy loss ΔE during a period of revolution around the nucleus is $\Delta E = -PT$, where $T = 2\pi r/v$. The speed of the electron on its orbit is $v = \sqrt{ar} = e/\sqrt{4\pi\epsilon_0 m_e r}$, so that

$$\Delta E = - \left(\frac{\mu_0}{3c} \right) \frac{e^5}{(4\pi\epsilon_0 m_e r)^{3/2} r}$$

Evaluating numerically:

$$\Delta E = -7.05 \times 10^{-24} \text{ J} = -4.4 \times 10^{-5} \text{ eV}$$

- b. The total energy of the electron is given by $E = -\frac{1}{2} \frac{e^2}{4\pi\epsilon_0 r_B} = -13.6 \text{ eV}$, then:

$$\frac{\Delta E}{E} = \frac{2\mu_0 e^3}{3c} \frac{1}{\sqrt{4\pi\epsilon_0} (m_e r_B)^{3/2}} = 3.24 \times 10^{-6}.$$

The energy lost during a period is negligible with respect to the total energy of the electron.

c. The emitted power corresponds to the rate of mechanical energy losses, i.e.,

$$\frac{dE}{dt} = -P.$$

If at instant t , the radius of the orbit is $r(t)$, we have:

$$E(t) = -\frac{1}{2} \left\{ \frac{e^2}{4\pi\epsilon_0 r(t)} \right\}.$$

Thus,

$$\frac{dE(t)}{dt} = \frac{dE(t)}{dr} \frac{dr}{dt} = \left(\frac{e^2}{8\pi\epsilon_0 r^2(t)} \right) \frac{dr}{dt} = -P,$$

leading to

$$\left(\frac{e^2}{8\pi\epsilon_0 r^2(t)} \right) \frac{dr}{dt} = - \left(\frac{\mu_0}{6\pi c} \right) \frac{e^6}{(4\pi\epsilon_0)^2 m_e^2 r^4(t)}$$

which is simplified into

$$\frac{dr(t)}{dt} = -\frac{4}{3} \frac{e^4}{(4\pi\epsilon_0)^2 m_e^2 c^3 r^2(t)} = -\frac{4}{3} c \frac{r_c^2}{r^2(t)},$$

where $r_c \equiv \frac{e^2}{4\pi\epsilon_0 mc^2} \frac{1}{mc^2}$ is the classical electron radius. Let τ be the time it takes the electron to collapse on the proton. Integrating the equation above by separating variables, we obtain

$$\int_{r_B}^0 r^2 dr = -\frac{4}{3} c r_c^2 \int_0^\tau dt,$$

where τ is the collapse time. We find

$$\frac{1}{3} r_B^3 = \frac{4}{3} r_c^2 c \tau$$

and

$$\tau = \frac{1}{4} \frac{r_B^3}{c r_c^2}$$

Evaluating the collapse time with $r_c = \alpha^2 r_B$ and $\alpha \approx 1/137$ yields

$$\tau \approx 1.6 \times 10^{-11} \text{ s.}$$

In conclusion, the classic view of the hydrogen atom consisting of an electron spinning around the proton is not consistent with electrodynamics: the latter predicts that a hydrogen atom does not survive more than 16 picoseconds!

Example 12.5 - Energy radiated by a charged particle

A particle of charge q falls radially on a repulsive potential $V(r) = \frac{\alpha}{r}$. The particle at infinity has a radial velocity v_0 . The particle radially approaches the origin up to a return point r_{\min} and then it goes back to infinity. Determine the total energy radiated by the particle during the entire journey.

Solution

The dipole moment of the charge q with respect to the origin is given by

$$\mathbf{p}(t) = qr(t)\mathbf{u}_r$$

Then

$$\ddot{\mathbf{p}}(t) = q \frac{d^2 r(t)}{dt^2} \mathbf{u}_r$$

The total power radiated by the charge at an instant t will be

$$P(t) = \left(\frac{\mu_0}{6\pi c} \right) q^2 \{ \ddot{r}(t) \}^2$$

Solving this problem exactly is really complex. However, we can assume that the total radiated energy is much smaller than the initial energy of the particle. In this way, we can assume that the trajectory is that of a particle that does not lose energy under the action of the potential. We can then determine r_{\min} satisfying

$$\frac{1}{2}mv_0^2 = \frac{q\alpha}{r_{\min}}.$$

Thus

$$r_{\min} = \frac{2q\alpha}{mv_0^2}.$$

Newton's equation also yields

$$m\ddot{r}(t) = -\frac{q\alpha}{r^2}$$

so that

$$P(t) = \left(\frac{\mu_0}{6\pi c} \right) \frac{q^4 \alpha^2}{m^2 r^4(t)}.$$

In addition, given the simplification of the problem, we can obtain the total radiated energy as twice the radiated energy on the path $r_{\min} \rightarrow \infty$. Thus

$$E = 2 \int_{t_0}^{t_f} dt P(t),$$

with

$$dt = \frac{dr}{v}$$

where

$$\frac{1}{2}mv^2 + \frac{q\alpha}{r} = \frac{1}{2}mv_0^2.$$

Then

$$v(r) = \sqrt{v_0^2 - \frac{2q\alpha}{mr}}.$$

Finally

$$\begin{aligned} E &= 2 \int_{r_{\min}}^{\infty} dr \left(\frac{\mu_0}{6\pi c} \right) \frac{q^4 \alpha^2}{m^2 r^4 \sqrt{v_0^2 - \frac{2q\alpha}{mr}}} \\ &= 2 \left(\frac{\mu_0 q^4 \alpha^2}{6\pi m^2 c} \right) \int_{r_{\min}}^{\infty} dr \frac{1}{r^4 \sqrt{v_0^2 - \frac{2q\alpha}{mr}}} \\ &= 2 \left(\frac{\mu_0 q^4 \alpha^2}{6\pi m^2 c} \right) \int_{r_{\min}}^{\infty} dr \frac{1}{r^4 \sqrt{\frac{2q\alpha}{m}} \sqrt{\frac{1}{r_{\min}} - \frac{1}{r}}} \\ &= 2 \left(\frac{\mu_0 q^4 \alpha^2}{6\pi m^2 c} \right) \sqrt{\frac{m}{2q\alpha}} \int_{r_{\min}}^{\infty} dr \frac{1}{r^4 \sqrt{\frac{1}{r_{\min}} - \frac{1}{r}}} \end{aligned}$$

This last integral can be evaluated numerically, leading to

$$E = 2 \left(\frac{\mu_0 q^4 \alpha^2}{6\pi m^2 c} \right) \sqrt{\frac{m}{2q\alpha}} \frac{16}{15 r_{\min}^{5/2}}.$$

Replacing the value of r_{\min} ,

$$E = \left(\frac{\mu_0 q m}{3\pi \alpha c} \right) \frac{2v_0^5}{15}.$$

Note that this result is valid while

$$E \ll \frac{1}{2} m v_0^2.$$

Example 12.6 - Electron in a gravitational field

An electron is released from rest and falls under the influence of gravity. What portion of the potential energy is radiated during the first centimeter?

Solution

The acceleration of the electron in the terrestrial gravitational field is equal to $a = g = 9.81 \text{ m/s}^2$, so that the radiated power is:

$$P = \left(\frac{\mu_0 e^2}{6\pi c} \right) g^2.$$

The time T that the electron takes to fall a distance d is given by

$$d = \frac{gT^2}{2}, \quad \text{i.e.,} \quad T = \sqrt{\frac{2d}{g}}$$

so that the energy radiated over a distance d is:

$$E_{\text{rad}} = PT = \left(\frac{\mu_0 e^2}{6\pi c} \right) g^2 \sqrt{\frac{2d}{g}}.$$

If $d = 1$ cm, we find $T = 4.51 \times 10^{-2}$ s, and

$$E_{\text{rad}} = 5.49 \times 10^{-52} \text{ J} = 3.42 \times 10^{-33} \text{ eV}.$$

On the other hand, the loss of potential energy after falling a distance d is:

$$\Delta U = -mgd = -8.94 \times 10^{-32} \text{ J} = -5.6 \times 10^{-13} \text{ eV}.$$

We see that the fraction of potential energy that is transferred in the form of radiation is negligible compared to ΔU :

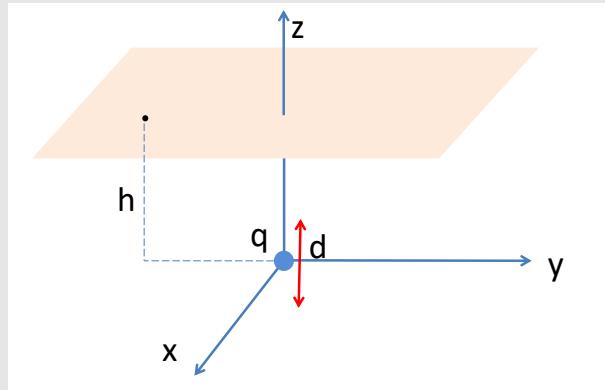
$$\frac{E_{\text{rad}}}{\Delta U} \sim 10^{-22}.$$

Then, virtually all the potential energy is converted into kinetic energy.

Example 12.7 - Oscillating charge

A charge q is initially at a distance h from an infinite plane. The charge begins to oscillate vertically around its initial position. Its position at time t satisfies $z(t) = d \cos \omega t$, where $h \gg d$.

- a. Determine the average power radiated through the plane, per unit area. Use appropriate coordinates with respect to the symmetry of the problem to represent the area element dS .
- b. Integrate the total power radiated through the entire plane, and compare with the total power radiated by the charge. Does this result make sense?



Solution

- a. Although the distance $H(t)$ between the oscillating charge and the plane varies in time according to

$$H(t) = h - d \cos \omega t,$$

since $d \ll h$, we can consider this distance as approximately constant, and equal to h . This is equivalent to locating the charge in a region very close to the origin of coordinates and then, sufficiently far from the origin ($|\mathbf{x}| = r \gg d$), the Poynting vector is given by

$$\Pi(\mathbf{x}, t) = \frac{\mu_0}{16\pi c^2 r^2} |\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c)|^2 \mathbf{u}_r$$

and the radiated power is given by

$$P = \frac{\mu_0}{6\pi c} |\ddot{\mathbf{p}}(t - r/c)|^2.$$

The dipole moment of the charge distribution is

$$\mathbf{p}(t) = qd \cos \omega t \mathbf{u}_z,$$

then

$$\ddot{\mathbf{p}}(t) = -qd\omega^2 \cos \omega t \mathbf{u}_z.$$

Thus, the total power radiated by the charge is

$$P = \left(\frac{\mu_0}{6\pi c} \right) q^2 d^2 \omega^4 \cos^2 (\omega(t - r/c)).$$

On the average, the charge radiates a power given by

$$\langle P \rangle = \left(\frac{\mu_0 q^2 d^2 \omega^4}{12\pi c} \right).$$

In addition, the Poynting vector takes the form

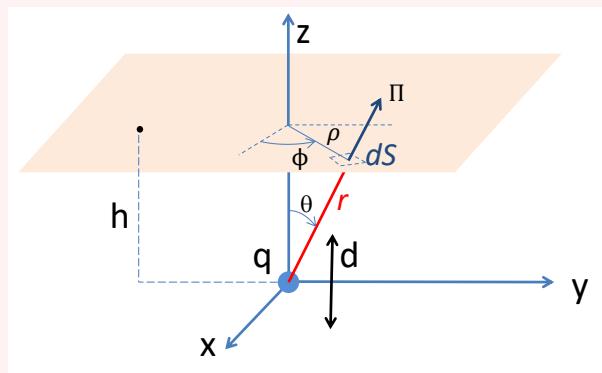
$$\begin{aligned} \Pi(\mathbf{x}, t) &= \frac{\mu_0}{16\pi^2 c r^2} \left| \mathbf{u}_r \times (-qd\omega^2 \cos \omega \tau \mathbf{u}_z) \right|^2 \mathbf{u}_r \\ &= \frac{\mu_0 q^2 d^2 \omega^4 \cos^2 \omega \tau}{16\pi^2 c r^2} \sin^2 \vartheta \mathbf{u}_r, \end{aligned}$$

with $\tau \equiv t - \frac{r}{c}$.

Averaging in time, we find

$$\langle \Pi(\mathbf{x}, t) \rangle = \frac{\mu_0 q^2 d^2 \omega^4}{32\pi^2 c r^2} \sin^2 \vartheta \mathbf{u}_r.$$

Now, let's calculate the power radiated through the plane per unit area. We will use the polar coordinates ρ and ϕ of the plane. An area element at the position determined by the coordinates (ρ, ϕ) is $dS = \rho d\rho d\phi$.



The power radiated (on the average) through the infinitesimal area element dS is

$$\begin{aligned} dP &= dS \mathbf{u}_z \cdot \langle \mathbf{\Pi} \rangle = dS |\langle \mathbf{\Pi} \rangle| \cos \vartheta \\ &= dS \frac{\mu_0 q^2 d^2 \omega^4}{32\pi^2 c r^2} \sin^2 \vartheta \cos \vartheta. \end{aligned}$$

Then, the power radiated through the plane per unit area is determined by integrating

$$\frac{dP}{dS} = \frac{\mu_0 q^2 d^2 \omega^4}{32\pi^2 c r^2} \sin^2 \vartheta \cos \vartheta.$$

The change of variables $r = \sqrt{h^2 + \rho^2}$, $\sin \vartheta = \frac{\rho}{\sqrt{h^2 + \rho^2}}$, $\cos \vartheta = \frac{h}{\sqrt{h^2 + \rho^2}}$, leads to

$$\frac{dP}{dS} = \frac{\mu_0 (qd\omega^2)^2}{32\pi^2 c} \frac{\rho^2 h}{(h^2 + \rho^2)^{5/2}}.$$

b. The total power (on average) through the plane will be

$$\begin{aligned} \langle P \rangle &= \int_0^\infty d\rho \int_0^{2\pi} d\phi \frac{\mu_0 (qd\omega^2)^2}{32\pi^2 c} \frac{\rho^3 h}{(h^2 + \rho^2)^{5/2}} \\ &= \frac{\mu_0 (qd\omega^2)^2 h}{16\pi c} \underbrace{\int_0^\infty d\rho \frac{\rho^3}{(h^2 + \rho^2)^{5/2}}}_{\frac{2}{3h}} \quad \rightarrow \quad \langle P \rangle = \frac{\mu_0 (qd\omega^2)^2}{24\pi c} \end{aligned}$$

We determined previously the total average power radiated by the charge

$$\langle P_{\text{total}} \rangle = \frac{\mu_0 (qd\omega^2)^2}{12\pi c}.$$

Naturally, the power radiated through the plane is half the total radiated power.

A way to obtain the result without performing the integration in the plane is to note that the Poynting vector is divergence free ($\nabla \cdot \mathbf{\Pi} = 0$), hence, the flux of the Poynting vector through a closed surface is zero. This means that average radiated power, or flux of the Poynting vector, through a disk of radius ρ centered on the z -axis in the plane at $z = h$ is equal to the flux of $\mathbf{\Pi}$ through the spherical cap bordered by the disk (center of the sphere at the origin and radius r), that is,

$$\langle P_{\vartheta_{\text{cap}}} \rangle = \int_0^{\vartheta_{\text{cap}}} \langle \mathbf{\Pi}(\mathbf{x}, t) \rangle \cdot \mathbf{u}_r r^2 2\pi \sin \vartheta d\vartheta = \frac{\mu_0 q^2 d^2 \omega^4}{32c\pi^2} 2\pi \int_0^{\vartheta_{\text{cap}}} \sin^3 \vartheta d\vartheta.$$

For the power radiated through the entire plane, $\vartheta_{\text{cap}} \rightarrow \pi/2$, and using $\int_0^{\pi/2} \sin^3 \vartheta d\vartheta = 2/3$, we retrieve

$$\langle P \rangle = \frac{\mu_0 (qd\omega^2)^2}{24c\pi}.$$

Example 12.8 - Charge suspended by a spring

Suppose that a charge of mass m and charge q is suspended from a spring with constant K . If the system is set to oscillate, how long must it take for the system to lose half of its initial mechanical energy?

Solution

Assuming that the charge oscillates in the z direction, the equation of motion, without taking into account any friction due to radiation, is given by:

$$m \frac{d^2 z}{dt^2} = -K(z - z_0),$$

where z_0 is the equilibrium position. The solution is a harmonic motion of amplitude A_0 around z_0 :

$$z(t) - z_0 = A_0 \cos(\omega t)$$

with $\omega = \sqrt{K/m}$. Assuming that the energy radiated during a cycle is much smaller than the mechanical energy of the system, we can consider that the acceleration corresponds at all times to $\mathbf{A}(t) = -A_0\omega^2 \cos(\omega t)\mathbf{u}_z$ and then the average power radiated during a cycle is:

$$\langle P \rangle = \left(\frac{\mu_0 e^2}{6\pi c} \right) \langle a^2 \rangle = \left(\frac{\mu_0 e^2}{6\pi c} \right) \frac{\omega^4 A_0^2}{2} = -\frac{dE}{dt}.$$

The mechanical energy of the system is given by $E(t) = \frac{1}{2}KA^2(t) = \frac{1}{2}m\omega^2 A^2(t)$, then:

$$\langle P \rangle = \left(\frac{\mu_0 \omega^2 e^2}{6\pi cm} \right) E(t) = -\frac{dE(t)}{dt}$$

so that the mechanical energy decays exponentially as $E(t) = E_0 e^{-t/\tau}$, with a time constant given by:

$$\tau = \frac{6\pi mc}{\mu_0 \omega^2 e^2}.$$

Finally, the mechanical energy will be equal to half of its initial value for $t = t^*$ such that:

$$e^{-t^*/\tau} = \frac{1}{2} \quad \rightarrow t^* = \tau \ln(2) = \frac{6\pi mc}{\mu_0 \omega^2 e^2} \ln(2).$$

Example 12.9 - Cyclotron radiation

Consider a charge q of mass m and initial velocity v_0 that describes a circular orbit in a uniform magnetic field B , perpendicular to the movement of the charge. Show that the radius R of the orbit evolves according to $R(t) = R_0 e^{-t/\tau}$. Calculate the total energy radiated by the particle.

Solution

The radius of the orbit is given by the balance between the acceleration and the Laplace force:

$$m \frac{v^2}{R} = qvB \rightarrow R = \frac{mv}{qB}.$$

Assuming that the radius of the orbit is approximately constant during a revolution, we have $a = v^2/R$ and the power radiated during a revolution is:

$$\langle P \rangle = \left(\frac{\mu_0 q^2}{6\pi c} \right) \frac{v^4}{R^2} = -\frac{dE}{dt}.$$

The kinetic energy at time t is given by

$$E(t) = \frac{1}{2}mv(t)^2 = \frac{1}{2m}R(t)^2q^2B^2,$$

so that:

$$\frac{dE}{dt} = R \frac{dR}{dt} \frac{q^2 B^2}{m}.$$

Then

$$\begin{aligned} -\frac{dE}{dt} &= -\frac{q^2 B^2}{m} R \frac{dR}{dt} = \left(\frac{\mu_0 q^2}{6\pi c} \right) \frac{v^4}{R^2} \\ &= \left(\frac{\mu_0 e^2}{6\pi c} \right) \frac{R^4 q^4 B^4}{m^4} \end{aligned}$$

Finally

$$\frac{dR}{dt} = -\left(\frac{q^4 B^2 \mu_0}{6\pi c m^3} \right) R$$

the solution of which is

$$R(t) = R_0 e^{-t/\tau}$$

with

$$\tau = \frac{6\pi c m^3}{q^4 B^2 \mu_0}.$$

As the radiated power can be written as a function of the radius as

$$P = \left(\frac{q^6 B^4 \mu_0}{6\pi c m^4} \right) R^2 = \left(\frac{q^6 B^4 \mu_0}{6\pi c m^4} \right) R_0^2 e^{-2t/\tau},$$

the total radiated energy is given by:

$$\Delta E = \left(\frac{q^6 B^4 \mu_0}{6\pi c m^4} \right) R_0^2 \int_0^\infty e^{-2t/\tau} dt = \left(\frac{q^6 B^4 \mu_0}{6\pi c m^4} \right) R_0^2 \frac{\tau}{2}$$

Introducing the expression for τ in the latter equation leads to

$$\Delta E = \left(\frac{q^2 B^2}{2m} \right) R_0^2 = \frac{1}{2} m v_0^2$$

which corresponds to the initial kinetic energy of the charge q .

12.7 Summary and essential formulas

In this chapter, all essential formulas were highlighted in definitions.

Section 12.3.4 summarizes the essential ingredients from which all other formulas are derived.

- For a source of electromagnetic radiation characterized by its charge density $\rho(\mathbf{r}, t)$ and current density $\vec{J}(\mathbf{r}, t)$, the scalar and the vector potentials read

$$V(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \iiint_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

- In the far zone

$$V_{\text{far}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0 r} \iiint_{\Omega} d^3x' \rho(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c).$$

$$\mathbf{A}_{\text{far}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\mathbb{R}^3} \mathbf{J}(\mathbf{x}', t - r/c + \mathbf{u}_r \cdot \mathbf{x}'/c) d^3x'.$$

- Electromagnetic radiation denotes the electromagnetic fields that transport energy arbitrarily far away from their source without attenuation of the total energy observed at distance r . The electromagnetic radiation field has the structure of a spherical wave with amplitudes of the electric and magnetic fields varying as $1/r$. Far from the source, \mathbf{E}_{rad} and \mathbf{B}_{rad} become quasi-plane waves propagating along \mathbf{u}_r , with \mathbf{E}_{rad} , \mathbf{B}_{rad} and \mathbf{u}_r forming a right-handed orthogonal triad that satisfies

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = \frac{\mathbf{u}_r}{c} \times \mathbf{E}_{\text{rad}}(\mathbf{x}, t).$$

- The radiated fields are entirely determined by the radiation vector potential

$$\mathbf{A}_{\text{rad}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*)$$

with

$$t^* = t - \frac{r}{c} + \frac{\mathbf{u}_r \cdot \mathbf{x}'}{c}$$

- The magnetic radiation field, reads

$$\mathbf{B}_{\text{rad}}(\mathbf{x}, t) = -\frac{\mathbf{u}_r}{c} \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t}$$

and the electric radiation field reads

$$\mathbf{E}_{\text{rad}}(\mathbf{x}, t) = \mathbf{u}_r \times \left(\mathbf{u}_r \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right)$$

- The electromagnetic power radiated through a spherical surface of radius r reads

$$P = \frac{1}{\mu_0 c} \int_0^{4\pi} d\Omega \left| \mathbf{r} \times \frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} \right|^2$$

where $d\Omega = \sin \vartheta d\vartheta d\phi$ denotes the infinitesimal element of solid angle defined with ϑ and ϕ , the polar and azimuthal angles of spherical coordinates.

- The *radiation vector*

$$\boldsymbol{\alpha}(\mathbf{x}, t) = \frac{\partial}{\partial t} \iiint_{\Omega} d^3x' \mathbf{J}(\mathbf{x}', t^*)$$

yields the expression for the vector potential

$$\frac{\partial \mathbf{A}_{\text{rad}}(\mathbf{x}, t)}{\partial t} = \frac{\mu_0}{4\pi r} \boldsymbol{\alpha}(\mathbf{x}, t),$$

which in turn leads to the angular distribution of radiation power,

$$\frac{dP}{d\Omega} = \frac{\mu_0}{16\pi^2 c} |\mathbf{u}_r \times \boldsymbol{\alpha}(\mathbf{x}, t)|^2.$$

- The total radiated power is obtained by integration over the solid angle:

$$P = \frac{\mu_0}{16\pi^2 c} \int_0^{4\pi} d\Omega |\mathbf{u}_r \times \boldsymbol{\alpha}(\mathbf{x}, t)|^2.$$

At large distances $r \rightarrow \infty$, the radiation vector $\boldsymbol{\alpha}(\mathbf{x}, t)$ and $\frac{dP}{d\Omega}$ only depend on the spherical angles θ and ϕ . The characterization of the radiation pattern consists in plotting the time average value for $\langle \frac{dP}{d\Omega} \rangle$ as a function of the polar and azimuthal angles ϑ and ϕ . The pattern exhibits lobes that provide information about the directivity of the source.

- For electric dipole radiation, the source is characterized by a time dependent dipole moment $\mathbf{p}(t)$ that is responsible for an electromagnetic radiation at the observation point \mathbf{x} and time t whose potential, vector potential, and electromagnetic fields are expressed as functions of the

retarded time $\tau = t - r/c$ as

$$V_{\text{rad,e}}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0 c r} \mathbf{u}_r \cdot \dot{\mathbf{p}}(t - r/c),$$

$$\mathbf{A}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \dot{\mathbf{p}}(t - r/c),$$

$$\mathbf{E}_{\text{rad,e}}(\mathbf{x}, t) = \frac{\mu_0}{4\pi r} \mathbf{u}_r \times (\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c)) = -\mathbf{u}_r \times c \mathbf{B}_{\text{rad,e}}(\mathbf{x}, t),$$

$$\mathbf{B}_{\text{rad,e}}(\mathbf{x}, t) = -\frac{\mu_0}{4\pi c r} (\mathbf{u}_r \times \ddot{\mathbf{p}}(t - r/c))$$

- The angular distribution of the radiation pattern for electric dipole radiation is given by

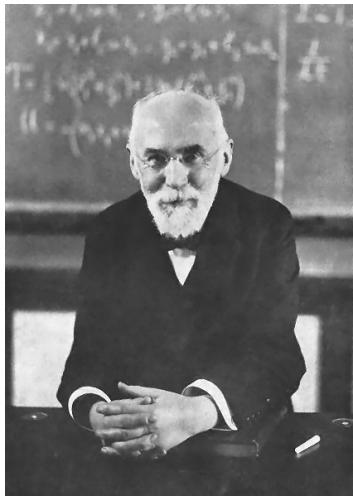
$$\left[\frac{dP}{d\Omega} \right]_e = \frac{\mu_0}{16\pi^2 c} |\ddot{\mathbf{p}}(t - r/c)|^2 \sin^2 \vartheta,$$

and is shaped as a doughnut with the symmetry axis (z -axis of spherical coordinates) aligned with the dipole moment. The total radiated power is expressed as

$$P = \frac{\mu_0}{6\pi c} |\ddot{\mathbf{p}}(t - r/c)|^2$$

Chapter 13

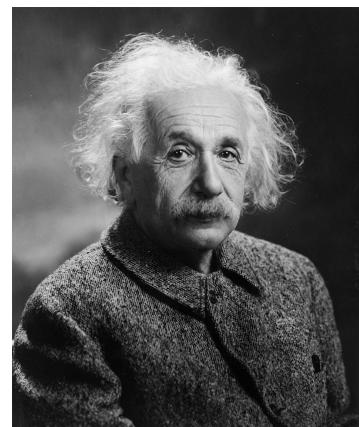
Special relativity



Lorentz



Minkowski



Einstein

13.1 Introduction

This chapter plays the role of a prerequisite chapter for the last part of this course of electrodynamics is devoted to the covariant formulation of Maxwell's equations, that is, a formulation that is manifestly invariant under a Lorentz transformation, which is the transformation between two Galilean frames, in the formalism of special relativity using rectilinear inertial coordinate systems. Before presenting this formulation, which will be the objective of the next chapter, we will therefore review in this chapter the theory of special relativity proposed by Einstein in 1905.

This chapter ends with an introduction to Minkowski's four dimensional space and associated four-vectors, which is required for the covariant formulation of Maxwell's equation. A brief presentation of the dynamics of relativistic particles serves both as an illustration of the use of four-vectors and as an application of Einstein's special relativity theory.

Alice and Bob will be recurrent fictional characters in this lecture. Although their characters were invented in 1978, well before the development of special relativity by Einstein, to illustrate thought experiments in cryptology, their numerous adventures in exchanging all types of messages makes them perfect archetypes agents for participating to most experiments of this course. In figures, point A will thus represent Alice while B will represent Bob. We will also need additional observers, Oliver (O) in the laboratory frame \mathcal{R} , and Oscar (O') in a moving frame \mathcal{R}' .

13.2 The relativity principle

13.2.1 Galileo's relativity principle in Newtonian mechanics

Until the nineteenth century, Galilean relativity treated all observers as equivalent no matter how fast they were moving. Consider the motion of a ball thrown by Alice straight up in the air at the North Pole; it falls straight back down again. Now Bob is doing the same experiment at the equator and observes that the same thing happens, even though he is moving at almost a thousand miles an hour faster than Alice at the pole.

Galileo's relativity principle states that the laws of physics remain the same in different reference frames, whether the frame is at rest or moving uniformly in a straight line with respect to a rest frame.

In order to underline what special relativity has changed, we briefly review useful definitions and assumptions of Newtonian mechanics.

- **SPACE-TIME FRAMEWORK IN NEWTONIAN MECHANICS**

A reference frame \mathcal{R} denotes an oriented system of coordinates, for instance Cartesian coordinates, equipped with rulers and clocks to measure position and time. The space-time framework of Newtonian mechanics then relies on the following assumptions and properties:

- An *event* is a localized phenomenon in time and space. It is given by its space-time coordinates in (x, y, z, t) a reference frame \mathcal{R} .

- The *physical space* is assumed to be the Euclidian space with three dimensions. It is homogeneous and isotropic, which means that no modification of physical properties can occur by translation or rotation of coordinate axes.

In an Euclidian space, the distance l between two fixed points $\mathbf{x}_1 \equiv (x_1, y_1, z_1)$ and $\mathbf{x}_2 \equiv (x_2, y_2, z_2)$ is obtained by the Euclidian distance

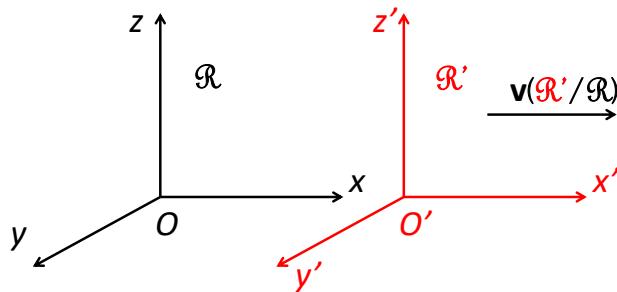
$$l^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

The distance l is invariant: it is the same for all observers. If Alice is at rest and holds a ruler of length l , Bob who looking at Alice while riding his bike at velocity \mathbf{v} , will see a ruler of the same length.

- Time* is defined by a reproducible phenomenon, that is, a clock, and represented by a parameter t . Time is absolute, meaning that it is independent of the reference frame. The duration Δt between two events is an invariant. With their respective clocks, Alice and Bob will measure the same duration for the ball motion from the instant Alice releases the ball to the instant Bob catches it while still his bike (Alice's clock is at rest and Bob's clock is in motion). The same events, for instance Bob catching the ball, take place at the same dates in different reference frames.

• GALILEAN TRANSFORMATION

In Newtonian mechanics, *Galilean* (or *inertial*) frames play a special role. Galilean frames are defined as the frames where objects move at constant velocity if they are not acted upon by external forces. The Galilean transformation is one of the most important concepts of Newtonian mechanics as it allows us to transform the space-time coordinates of an event in a Galilean frame \mathcal{R} into the space-time coordinates of the same event viewed from another Galilean frame \mathcal{R}' , moving at uniform velocity \mathbf{v} with respect to \mathcal{R} .



For instance, assume two Galilean frames $\mathcal{R}(0, x, y, z)$ and $\mathcal{R}'(0', x', y', z')$ such that \mathcal{R}' moves at constant velocity \mathbf{v} with respect to \mathcal{R}' ($\mathbf{v} \equiv \mathbf{v}(\mathcal{R}/\mathcal{R}')$). The coordinate axes can always be chosen in such a way that the x -axis is parallel to the x' -axis and to \mathbf{v} . The Galilean transformation from \mathcal{R} to \mathcal{R}' reads

$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \\ t' = t \end{cases}$$

and the inverse transformation reads

$$\begin{cases} x = x' + vt' \\ y = y' \\ z = z' \\ t = t' \end{cases}$$

The inverse transformation is simply obtained by changing the sign of the velocity, since from the point of view of an observer at rest in \mathcal{R}' , it is the reference frame \mathcal{R} that is moving at velocity $-\mathbf{v}$ with respect to \mathcal{R}' .

What are the implications of the Galilean transformation on velocities? Suppose Alice, observer in \mathcal{R} , throws a ball along the x -axis. Alice observes the ball moving at velocity $\mathbf{v}_{\text{ball}}(t) = \frac{dx(t)}{dt} \mathbf{u}_x$. Bob, observer from his bike in \mathcal{R}' will see the ball moving at velocity $\mathbf{v}'_{\text{ball}} = \frac{dx'(t')}{dt'} \mathbf{u}'_x$. Differentiating the coordinates of the ball with respect to time in each frame leads to the velocities observed by each observer:

$$\frac{dx'(t')}{dt'} = \frac{dx'}{dt} \underbrace{\frac{dt}{dt'}}_1 = \frac{d(x(t) - vt)}{dt} = \frac{dx(t)}{dt} - v$$

Thus,

$$v'_{\text{ball}} = v_{\text{ball}} - v.$$

This is the composition law for velocities.

Differentiating again with respect to time yields accelerations:

$$a'(t') = \frac{d}{dt'} \mathbf{v}'_{\text{ball}} = \underbrace{\frac{dt'}{dt}}_1 \left(\frac{d}{dt} \mathbf{v}_{\text{ball}} - \underbrace{\frac{dv}{dt}}_0 \right) = \frac{d}{dt} \mathbf{v}_{\text{ball}} = a(t).$$

Hence Alice and Bob observe the same acceleration of the ball. They conclude that Newton's law

$$\mathbf{F} = m\mathbf{a}$$

is the same in both reference frames.

Definition 13.1: Galileo's relativity principle

- The laws of mechanics are the same in all Galilean reference frames. Galileo's transformation leaves invariant the Newton laws.
- These laws are said to be covariant with respect to a change of Galilean reference frame.

Question: Can we extend this relativity principle to the laws of electrodynamics, while preserving the same space-time framework?

13.2.2 Electromagnetism and relativity principle

• MAXWELL EQUATIONS (1854)

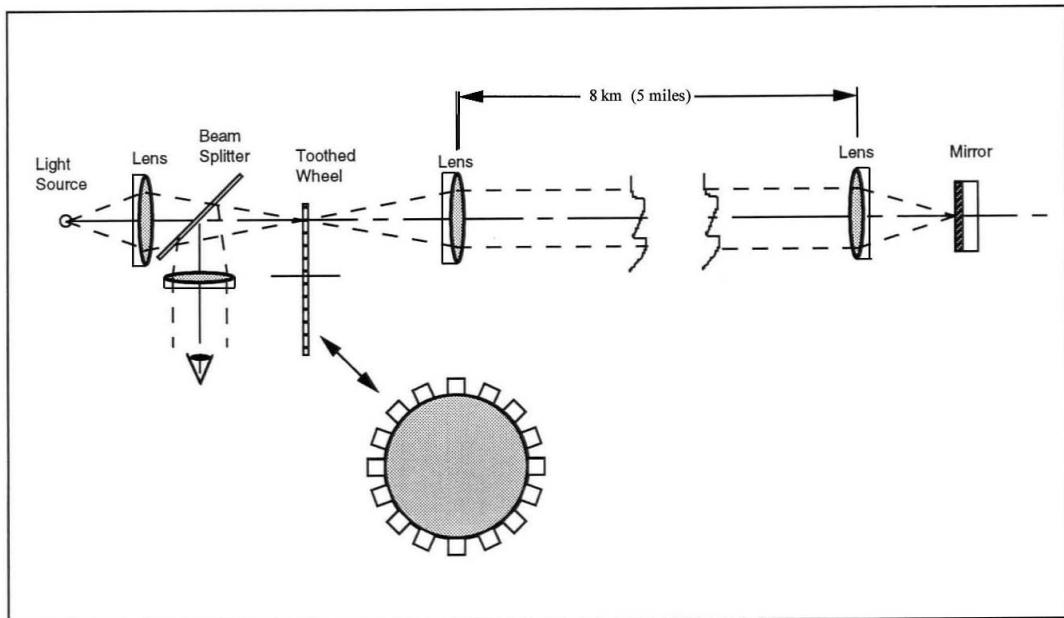
Maxwell's theory, published in 1854, established that at a given instant and at each point of space, properties of space are modified by a *charge distribution in motion*. The charge distribution is responsible for the generation of an electromagnetic field, characterized by an electric field **E** and a magnetic field **B**.

A charge q undergoes the action of the electromagnetic field via the force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

The electromagnetic field, **E** and **B**, can be calculated from Maxwell's equations. As we learned throughout the course, Faraday showed that there is a second source for electric fields: a change in time of a magnetic field is also the source of an electric field (induction phenomenon). Maxwell then postulated that a change in time of an electric field is also a source of magnetic field. This led to the prediction that electromagnetic fields propagate in vacuum at velocity $v = (\epsilon_0 \mu_0)^{-1/2}$, where μ_0 and ϵ_0 characterize the electric properties of vacuum. Up to measurement uncertainties, it was shown that $v \sim c$, that is, electromagnetic waves propagate at the velocity of light. This led to the conclusion that light is an electromagnetic wave.

Fizeau was the first to succeed in a terrestrial measurement of the velocity of light in 1849, sending a light beam along a 17 km round-trip path across the outskirts of Paris. At the light source, the exiting beam was chopped by a rotating toothed wheel; the measured rotational rate of the wheel at which the beam, upon its return, was eclipsed by the toothed rim was used to determine the travel time of the beam. Fizeau reported a light speed that differs by only about 5 percent from the currently accepted value.



FIZEAU'S EXPERIMENT

Hertz produced radio waves for the first time in 1888 and confirmed Maxwell's prediction that they propagate at the same velocity as light, and undergo similar phenomena of reflection, refraction and polarization.

• MAXWELL EQUATIONS AND RELATIVITY PRINCIPLE

At the end of the XIXth century, scientists knew that electromagnetic waves propagate at velocity $v = (\epsilon_0 \mu_0)^{-1/2}$. If the velocity of electromagnetic waves predicted by Maxwell's equations were only valid in a specific reference frame, this frame would have an absolute nature, that is, Maxwell's equations would not be valid in any reference frame. Experiments on light waves would allow to single out this specific frame. Two point of views were then debated to conciliate Maxwell's theory and the concept of absolute reference frame originating from classical Newtonian mechanics.

- Scientists in the XIXth century have assumed that Maxwell's equations needed modifications and concluded that a medium was necessary for the propagation of electromagnetic waves and light, and that the speed of light with respect to that medium was constant, just like the propagation of sound in air. This medium was named *Aether* and was seen as an invisible, massless, jelly-like medium filling all space and fixed with respect to an absolute reference frame identified as the Copernic reference frame. Aether had other exotic properties like its high rigidity, allowing light to travel so fast in it while allowing Earth or objects to move in it without any drag. Needless to say that no consensus was reached regarding Aether.
- There was a second point of view: Maxwell's equations are equally valid in *all* Galilean reference frames, that is, they obey the relativity principle. As a result, the velocity of electromagnetic waves is unique: it is the same in all reference frames. The main problem is that this principle is not compatible with the laws of Newtonian mechanics. If we accept this principle, then,

The space-time framework of physics is not that of classical kinematics.

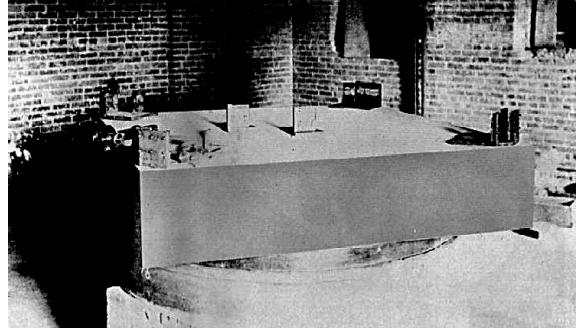
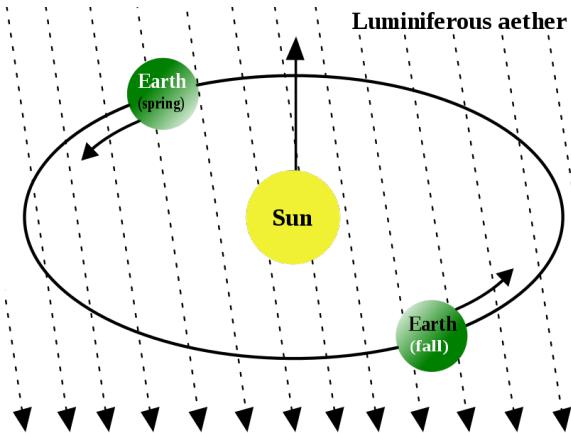
Experiments have proven this second point of view to be valid.

13.2.3 Experimental results

In this section, we review a couple of important experiments that allowed scientists to refute the concept of Aether on the one hand, and to confirm the validity of the second point of view on the other hand.

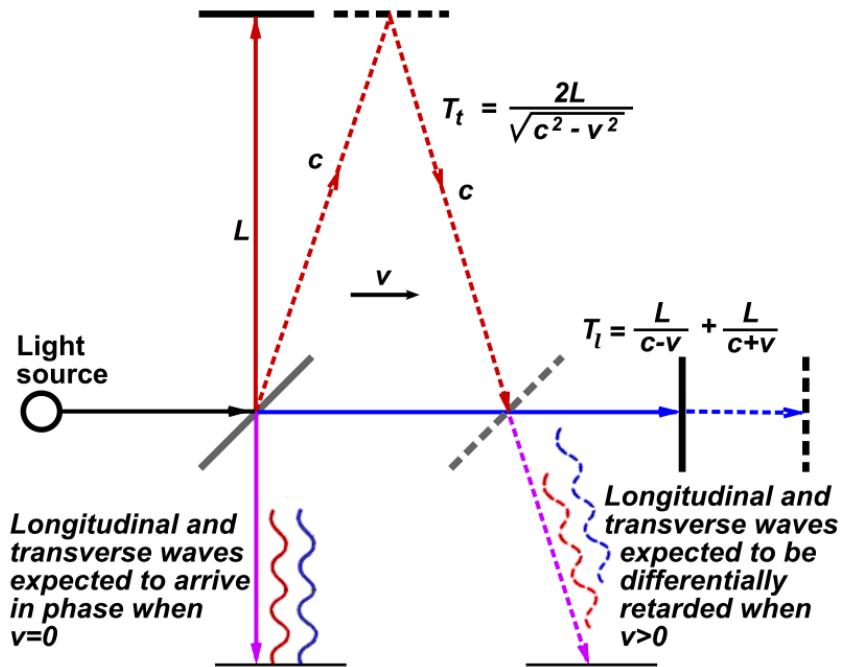
• MICHELSON AND MORLEY'S EXPERIMENT

Michelson and Morley attempted to detect the existence of aether in a famous experiment in 1887. Their goal was to measure the orbital velocity of the Earth by optical means. Assuming the Earth is moving through the aether, the speed of light in the direction of Earth's motion should be lower than it is in a perpendicular direction. Michelson and Morley thus built an experiment to measure these speeds in order to detect Earth's absolute velocity relative to the aether.



Michelson and Morley's interferometric setup

The principle of their experiment is the following: A light source sends light into the two arms of an interferometer, thanks to a glass plate that partly transmits the (blue ray) beam while it partly reflects the (red ray) beam (see continuous lines in the figure). In both arms, the light beams are reflected by mirrors and recombine after hitting the glass plate again, which is totally reflective on the side of the blue ray. The recombination leads to interference fringes that depend on the length difference between the two arms.



The basic assumption to understand the principle of the experiment is that the Earth constitute a frame $\mathcal{R}'_{\text{abs}}$ that moves in aether at the absolute velocity $\mathbf{v} \equiv \mathbf{v}_{\text{Earth}/\mathcal{R}'_{\text{abs}}}$. The light velocity on the Earth is then different from that in Aether $\mathbf{c} = \mathbf{c}_{\mathcal{R}'_{\text{abs}}}$, satisfying

$$\mathbf{c}_{\mathcal{R}'_{\text{abs}}} = \mathbf{c}'_{\mathcal{R}'_{\text{Earth}}} + \mathbf{v}_{\text{Earth}/\mathcal{R}'_{\text{abs}}}$$

Now, if one of the arms of the interferometer is parallel to the velocity of Earth in aether, there will be a difference of phase between the two arms, that is, the time required by light to make the round

trip from the glass plate to the mirror and back to the glass plate will differ for the longitudinal ray parallel to \mathbf{v} and for the perpendicular ray. For the longitudinal arm, the light indeed travels toward the mirror at velocity c during a time T_{l1} , over a distance equal to the sum of the length of the arm and the distance traveled by the mirror, hence $L + vT_{l1} = cT_{l1}$, leading to

$$T_{l1} = \frac{L}{c - v}$$

On its way back, the light will travel a shorter distance to hit the glass plate, $L - vT_{l2} = cT_{l2}$, leading to a travel time

$$T_{l2} = \frac{L}{c + v}$$

The total duration of the light journey in the longitudinal arm is then

$$T_l = T_{l1} + T_{l2} = \frac{2Lc}{c^2 - v^2}$$

In the perpendicular arm, the light travels at c along twice the length of the hypotenuse of a right angle triangle, hence the travel time satisfies

$$(cT_t/2)^2 = L^2 + (vT_t/2)^2 \quad \rightarrow \quad T_t = \frac{2L}{\sqrt{c^2 - v^2}}.$$

The time difference between the longitudinal and the transverse arms is at leading order in v/c :

$$(T_l - T_t) \sim \frac{L}{c} \frac{v^2}{c^2}$$

and is thus responsible for a phase difference $\Delta\phi = 2\pi c(T_l - T_t)/\lambda$ where λ is the wavelength of the light source. For the orbital velocity of the Earth, $v \sim 3 \times 10^4$ m/s, arms of length $L = 10$ m and a wavelength of 600 nm, we find

$$\Delta\phi \sim 2\pi \frac{L}{\lambda} \frac{v^2}{c^2} = 2 \times 3 \frac{10}{600 \times 10^{-9}} \frac{9 \times 10^{-8}}{9 \times 10^{-16}} = 1 \text{ rad},$$

which is large enough to be visible.

Now it is not guaranteed that the two arms of the interferometer have exactly the same length, thus leading to a fringe pattern even if $v = 0$. However, to disambiguate the fringes due to the imperfect setup from those due to the motion of Earth in Aether, it is sufficient to rotate the interferometer by 90 degrees, thus exchanging the roles of the arms, so as to observe a change in the fringe pattern corresponding to the Earth motion.

In spite of numerous experiments performed at different times of the day and year, the result was always negative. Michelson and Morley could not observe a change in the fringe pattern and concluded that if there is a motion of Earth with respect to Aether, the velocity must be extremely small. An absolute motion could not be singled out by means of electromagnetic waves.

• MEASUREMENT OF THE VELOCITY OF LIGHT

Measurement of the velocity of light produced by a moving source were performed by Alväger et al. in 1964, at the CERN Proton Synchrotron. They executed a time of flight measurement to test

the Newtonian momentum relations for light, being valid in the so-called emission theory. In this experiment, gamma rays were produced in the decay of 6-GeV pions π^0 traveling at $0.99975 c$. If the Newtonian momentum $p = mv$ were valid, those gamma rays should have traveled at superluminal speeds. However, testing a Newtonian relation in the form $c' = v + c$, they found no difference with a source at rest:

$$\text{Accelerated pion: } c'_{\text{Lab}} = (2.9977 \pm 4 \times 10^{-4}) \times 10^8 \text{ m/s}$$

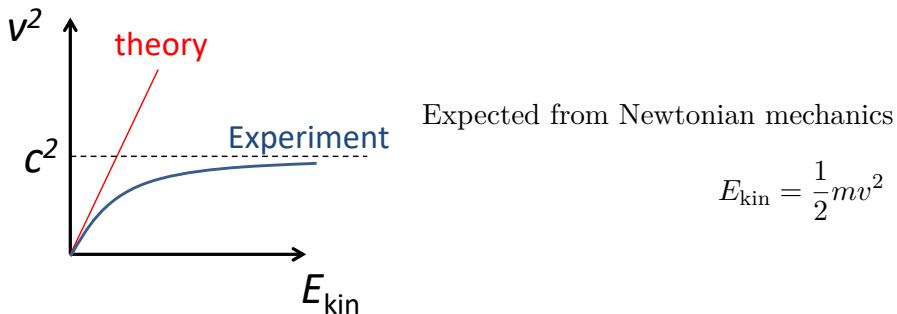
$$\text{Source at rest: } c_{\text{Lab}} = (2.99792458 \pm 1.2) \times 10^8 \text{ m/s}$$

Hence, they gave an upper limit of $\Delta v/c \sim 10^{-5}$.

They concluded that $c \simeq c'$: the velocity of light is independent of the reference frame. This result is not compatible with Newtonian mechanics.

• BERTOZZI'S EXPERIMENT (1964)

Bertozzi's experiment was performed in 1964. Special relativity (and general relativity) had already been successfully tested by earlier experiments but the relativistic dynamics of particles had been tested only indirectly. The development of particle accelerators in the XXth century allowed to perform measurements with increasing accuracy to test the predictions of relativity. Bertozzi's experiment is the first experiment leading to a direct measurement of the kinetic energy of accelerated particles. Using an electron accelerator facility at MIT, electrons were accelerated by very high electric fields and their velocity (energy) was measured. The experiment showed an upper velocity limit for accelerated particles. Newtonian mechanics simply predicts that the kinetic energy of electrons can increase up to any value, linearly with v^2 . In contrast, Bertozzi's experiments demonstrated a saturation. The velocity of the particle could not exceed the velocity of light.



In conclusion, all experiments that attempted to single out an absolute reference frame were negative. This turned out to be impossible. The light velocity in vacuum does not depend on the propagation direction. Space is isotropic. The light velocity does not depend on the reference frame.

Maxwell's equations satisfy the relativity principle. It is Newtonian mechanics that must be cured, not Maxwell's equations.

13.2.4 Einstein's special relativity principle (1905)

In 1905, Einstein published the theory of special relativity, which follows from the enunciation of the relativity principle.

Definition 13.2: Relativity principle

- **Postulate:** The fundamental laws of physics are covariant with respect to a change of Galilean reference frame
- The velocity of light in vacuum is invariant. Light propagates in vacuum isotropically at velocity c with respect to any Galilean reference frame.
- **Postulate:** A velocity limit exists for the propagation of a signal, corresponding to the velocity of light in vacuum.

A presentation of special relativity will be detailed in the next sections. We note that a requirement of the new theory is the equivalence principle, stating that Newtonian mechanics must be retrieved as a particular case of special relativity.

Definition 13.3: Equivalence principle

If $v \ll c$, the laws of Newtonian mechanics must be recovered.

For the purpose of showing that Maxwell's theory is fully compatible with the relativity principle, special relativity is all we need. However, special relativity is not the end of the story. In 1916, Einstein published another fundamental paper in the history of physics, presenting the theory of general relativity, which extends the theory of special relativity to the case of non Galilean reference frames. General relativity goes beyond the scope of this course.

13.3 Relativistic kinematics

Einstein's postulates mean that our Universe is different from the Newtonian Universe we are daily experiencing. In this Universe, velocities of material objects cannot be infinite and Newton's concept of absolute time does not hold. The simultaneity of events for all observers no longer holds as well. The kinematics characterizing this Universe without infinite velocity is presented in this section.

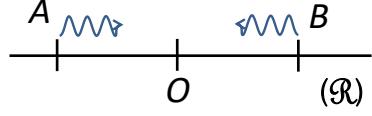
13.3.1 Relative nature of time

What is simultaneity? How to decide whether two events are simultaneous in a Universe without infinite velocity?

Two events are simultaneous if they happen at the same instant. Therefore, we need clocks to help us measure simultaneity. Since the two events we wish to measure may happen at different locations, we also need to synchronize the clocks used for the measurement.

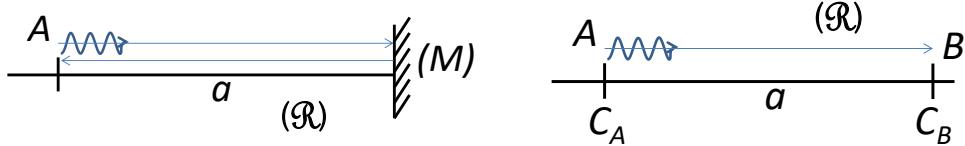
Suppose Alice and Bob are at fixed positions, at rest, in a reference frame \mathcal{R} . An observer, Oliver, stands exactly in the middle of Alice and Bob. Alice and Bob emit an electromagnetic signal toward

Oliver. We know that these signals travel at c , the velocity of light¹. If Oliver receives the signals from Alice and Bob at the same instant, he observes that the two events (emission of signal S_A by Alice and S_B by Bob) are simultaneous. It is then possible for Alice and Bob to synchronize their clocks.



- SYNCHRONIZATION OF FIXED CLOCKS IN A REFERENCE FRAME \mathcal{R}

To observe simultaneous events in a given reference frame is therefore possible but it requires to synchronize fixed clocks. Synchronization can be achieved in the following way: If the propagation velocity of the signal is unknown, it is possible to first measure it. To do so, Alice first emits a signal (at $t = 0$) toward Bob, who holds a mirror, a distance a away from her. She will receive the signal at $t = 2a/c$ and knowing the distance, she can infer the propagation velocity of the signal in \mathcal{R} .



Then Alice emits another signal at t_0 in Bob's direction. Bob will receive the signal at $t_0 + \frac{a}{c}$ and can thus synchronize his clock with that of Alice. Fixed clocks in a reference frame \mathcal{R} can be synchronized.

- RELATIVITY (NON ABSOLUTE NATURE) OF SIMULTANEOUS EVENTS

Now consider another observer in a different reference frame, Oscar, at O' in the frame \mathcal{R}' moving at velocity v with respect to \mathcal{R} . These frames are one-dimensional Universes with x -axis parallel to the velocity v . Oscar and Oliver stand at O' and O , respectively, and synchronize their clocks when they pass in front of each other. At this instant ($O' \equiv O$), $t = t' = 0$.



At the same instant, Alice and Bob who live in \mathcal{R} and have clocks that are synchronized with that of Oliver, emit their signals S_A (from A), and S_B (from B). Of course, nothing changed from the point of view of Oliver in reference frame \mathcal{R} . He receives the signals at the same instant. However, Oscar in \mathcal{R}' observes that Oliver does not receive the signals at the same time. From Oscar's point of view, the

¹If Alice and Bob use a slower communication means, for instance pigeons rather than radio waves, they will simply follow the same procedure, adding a preliminary measurement of the velocity c of the signal

signal from Alice travels at velocity $c - v$ while the signal from Bob travels at $c + v$. Oscar therefore concludes that the signals are out of synchronization by an amount

$$\Delta T = \frac{a/2}{c - v} - \frac{a/2}{c + v} = \frac{a}{c} \frac{v/c}{(1 - v^2/c^2)}$$

This is a very small number if $v \ll c$. Nevertheless, Oscar concludes that the events (reception of signals S_A and S_B by Oliver) are not simultaneous, while Oliver receives the signals at the same time. When v approaches c , the delay observed by Oscar even becomes arbitrarily large. Simultaneity depends on the reference frame. Time is relative to the observer.

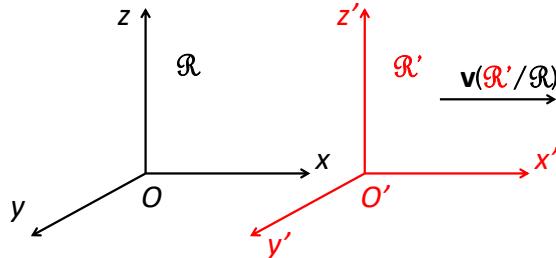
Consequently, the postulate of the absolute nature of time in Newtonian mechanics turns out to be an approximation of the kinematics in a Universe without infinite velocity; it breaks down when the relative velocity between frames approaches the speed of light.

The invariance of the velocity limit in the Universe means that the velocity of an electromagnetic signal is the same in \mathcal{R}' as in \mathcal{R} : $\mathbf{c}_{S_B/\mathcal{R}} = \mathbf{c}_{S_B/\mathcal{R}'}$, which implies that the space-time framework need to be reconstructed.

13.3.2 Lorentz's special transformation

Consider an event of space-time coordinates (x, y, z, t) in a Galilean frame \mathcal{R} . In another Galilean frame \mathcal{R}' , the same event has space-time coordinates (x', y', z', t') . The Lorentz transformation is the kinematic transformation that replaces the Galilean transformation of Newtonian mechanics, that is, which is linking space-time coordinates of an event in \mathcal{R} to those in \mathcal{R}' . The new transformation must be compatible with the relativity principle. It is demonstrated from homogeneity and isotropy of space, invariance of c and the relativity principle.

The two Galilean frames \mathcal{R} and \mathcal{R}' are thus equipped with rulers and clocks. All clocks of a given frame are synchronized. The frame \mathcal{R}' moves at velocity $\mathbf{v} = \mathbf{v}_{\mathcal{R}'/\mathcal{R}}$ with respect to \mathcal{R} .



The origins of \mathcal{R} and \mathcal{R}' are chosen so as to coincide when the clocks in \mathcal{R} mark $t = 0$. The clock attached to \mathcal{R}' placed at O' marks $t' = 0$ when O' is at O , that is, the origin events coincide in \mathcal{R} and \mathcal{R}' . We chose the axes of coordinates and their orientation in \mathcal{R} and \mathcal{R}' such that $\mathbf{v}_{\mathcal{R}'/\mathcal{R}} \parallel Ox \parallel O'x'$, and in the transverse direction, $Oy \parallel O'y'$, and $Oz \parallel O'z'$.

Before quoting the Lorentz transformation, we define standard notations in special relativity:

- The β factor represents a dimensionless velocity of the frame in motion, normalized by the light velocity. It is therefore smaller than one.

$$\beta = \frac{v}{c} \leq 1.$$

- The Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}} \geq 1$$

can take all values from 1 to infinity.

The Lorentz transformation and its inverse then reads

$$\begin{aligned}x' &= \frac{1}{\sqrt{1 - v^2/c^2}}(x - vt) \\y' &= y \\z' &= z \\t' &= \frac{1}{\sqrt{1 - v^2/c^2}} \left(t - \frac{vx}{c^2} \right)\end{aligned}$$

$$\begin{aligned}x &= \frac{1}{\sqrt{1 - v^2/c^2}}(x' + vt) \\y &= y' \\z &= z' \\t &= \frac{1}{\sqrt{1 - v^2/c^2}} \left(t' + \frac{vx'}{c^2} \right)\end{aligned}$$

or in terms of the β and Lorentz (γ) factors:

$$\begin{aligned}x' &= \gamma(x - \beta ct) \\y' &= y \\z' &= z \\ct' &= \gamma(ct - \beta x)\end{aligned}$$

$$\begin{aligned}x &= \gamma(x' + \beta ct) \\y &= y' \\z &= z' \\ct &= \gamma(ct' + \beta x')\end{aligned}$$

The Lorentz transformation ensures that the equivalence principle is satisfied, which means that if the velocity of \mathcal{R}' is much smaller than the velocity of light, the Lorentz transformation tends to the Galilean transformation

$$v \ll c \quad \rightarrow \quad \gamma \sim 1 \quad \rightarrow \quad \begin{cases} x' \sim x - vt, \\ t' \sim t. \end{cases}$$

Hence, the new kinematics is compatible with Newtonian mechanics.

13.3.3 Invariance of space-time intervals

The Lorentz transformation introduces a coupling between space and time. This has several consequences, among which a change in the invariant quantities, that is the quantities that take the same numerical value in all Galilean frames. For instance, the speed of light is an invariant quantity by Einstein's postulate. A second invariant quantity is the charge of a particle. The space-time interval constitute a third important invariant quantity, as it allows us to distinguish past events from future events and thus, causes from effects. If we consider two events, which have space-time coordinates in \mathcal{R} and \mathcal{R}' :

$$\begin{aligned}\text{In } \mathcal{R} : \quad &\text{Event 1 } (x_1, y_1, z_1, t_1); \quad \text{Event 2 } (x_2, y_2, z_2, t_2), \\ \text{In } \mathcal{R}' : \quad &\text{Event 1 } (x'_1, y'_1, z'_1, t'_1); \quad \text{Event 2 } (x'_2, y'_2, z'_2, t'_2),\end{aligned}$$

the square of the space time interval between the two events is defined as

$$\text{In } \mathcal{R} : \Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2$$

$$\text{In } \mathcal{R}' : \Delta s'^2 = \Delta x'^2 + \Delta y'^2 + \Delta z'^2 - c^2 \Delta t'^2$$

We will show that the space-time interval is invariant:

$$(\Delta s)^2 = (\Delta s')^2 = (\Delta l)^2 - c^2(\Delta t)^2.$$

Expanding the space time interval in \mathcal{R}' and replacing the event coordinates in \mathcal{R}' by those in \mathcal{R} using the Lorentz transformation and the fact that $\gamma^2(1 - \beta^2) = 1$, we have

$$\begin{aligned} \Delta s'^2 &= (x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - (ct'_2 - ct'_1)^2 \\ &= \gamma^2(x_2 - vt_2 - x_1 + vt_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - \gamma^2(ct_2 - \beta x_2 - ct_1 + \beta x_1)^2 \\ &= \gamma^2(x_2 - x_1)^2(1 - \beta^2) - \gamma^2(t_2 - t_1)^2c^2(1 - \beta^2) + \Delta y^2 + \Delta z^2 \\ &= (x_2 - x_1)^2 - c^2(t_2 - t_1)^2 + \Delta y^2 + \Delta z^2 \\ &= \Delta s^2 \end{aligned}$$

Therefore, we find

$$\Delta s = \pm \Delta s'$$

and the limit at low velocity allows us to choose the positive sign:

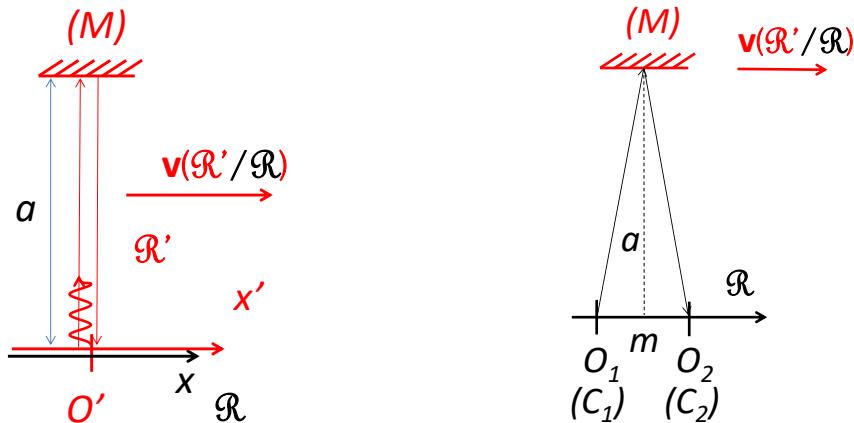
$$\text{if } v \rightarrow 0, \quad \Delta s' \rightarrow \Delta s \quad \Rightarrow \quad \Delta s = \Delta s'.$$

The space-time interval is an invariant.

13.3.4 Time dilation

- PROPER TIME

The proper time denotes the interval between two events taking place at the same point of a Galilean reference frame. It is therefore measured by a clock attached to the Galilean frame. This frame is called the proper reference frame for these two events.



For instance, consider two events taking place in a Galilean reference frame \mathcal{R}' moving at velocity v with respect to the frame \mathcal{R} : Event 1 corresponds to the emission of an electromagnetic signal from O' in the direction of the y -axis, and event 2 corresponds to the reception of the signal at O' after it is reflected by a mirror attached to \mathcal{R}' , a distance a away from O' . The two events take place at the same point O' of \mathcal{R}' , hence, the interval between the two events is the proper time interval in \mathcal{R}' . Now let O_1 be the point in \mathcal{R} that coincides with O' at event 1, when the signal is emitted. Similarly, O_2 denotes the point in \mathcal{R} that coincides with O' at event 2, when the signal is received. If we use a clock of \mathcal{R} to measure the time interval Δt between the two events viewed from \mathcal{R} , we find

$$\Delta t = 2 \frac{O_1 M}{c} \rightarrow O_1 M = \frac{c \Delta t}{2},$$

while the interval between the two events viewed from the proper frame \mathcal{R}' is

$$\Delta t' = \frac{2a}{c} \rightarrow a = \frac{c \Delta t'}{2}.$$

In the triangle $O_1 m M$, we can use Pythagoras's relation to link the time intervals measured by the two clocks in \mathcal{R} and in \mathcal{R}' :

$$(O_1 M)^2 = a^2 + (O_1 m)^2, \quad \text{with} \quad O_1 m = \frac{v \Delta t}{2}$$

Replacing all distances by their expressions as functions of the time intervals, we have

$$\frac{\Delta t^2 c^2}{4} = \frac{(\Delta t')^2 c}{4} + \frac{v^2 \Delta t^2}{4}$$

Thus,

$$\Delta t^2 (c^2 - v^2) = c^2 \Delta t'^2 \rightarrow \Delta t = \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}}$$

We find

$$\boxed{\Delta t = \gamma \Delta t'}$$

and since $\gamma \geq 1$, the proper time is smaller than the time interval measured in \mathcal{R} : $\Delta t' \leq \Delta t$.

The observer in the laboratory reports a longer elapsed time than the observer in the moving frame.

• TIME DILATION EQUATION

The expression for the elapsed time between two events observed by an observer who is not in the proper reference frame is called the time dilation equation. It can be obtained directly from the Lorentz transformation between the Galilean frames \mathcal{R} and \mathcal{R}' , moving at v with respect to \mathcal{R} .

Indeed if two events are localized at the same point in the Galilean frame \mathcal{R}' , they occur at times t'_1 and t'_2 , respectively, and at the same position $x'_1 = x'_2$, that is $\Delta x' = 0$. In the reference frame of the observer \mathcal{R} , we have from the Lorentz transformation

$$t = \gamma \left(t' + \frac{vx'}{c^2} \right)$$

The elapsed time for the observer thus reads

$$\Delta t = \gamma(\Delta t' + \frac{v\Delta x'}{c^2}) \quad \rightarrow \quad \boxed{\Delta t = \gamma\Delta t'}.$$

The observer in \mathcal{R} reports a longer elapsed time than the observer in the moving frame \mathcal{R}' . However, if two events occur in \mathcal{R} at the same point $x_1 = x_2$, then $\Delta x = 0$, and using the Lorentz transformation, we find, following similar arguments,

$$t' = \gamma \left(t - \frac{vx}{c^2} \right) \quad \rightarrow \quad \boxed{\Delta t' = \gamma\Delta t}$$

The observer in the moving frame \mathcal{R}' observes a longer elapsed time than the observer in \mathcal{R} . The situation seems to be perfectly symmetric. From the point of view of \mathcal{R}' , it is the reference frame \mathcal{R} that is moving at velocity $-v$.

This means that extreme care is required when applying the time dilation equation. It is important to appreciate the role of the proper reference frame in which two events occur at the same position.

The time dilation is not a weird consequence of the kinematics arising from the Lorentz transformation but a real phenomenon that was experimentally verified from the radioactive decay of muons (mu-mesons) from cosmic rays. These particles have a finite lifetime corresponding to the characteristic exponential decay time of an initial population. This lifetime is independent on the initial number of particles. The exponential decay then serves as a proper clock, measuring the proper time for the muons. The decay time experimentally was found to depend on the mean-velocity of muons in a way that agrees very well with the time dilation equation.

• PROPER TIME FOR A PARTICLE, PROPER REFERENCE FRAME

Even if it is accelerated, a particle can be considered as being at rest in a Galilean reference frame \mathcal{R}_0 having the particle velocity, at time t . This frame is called the tangent reference frame as its velocity with respect to the laboratory frame coincides with that of the particle at time t .

Consider the position of the particle at $M(x, y, z)$ and its velocity $\mathbf{v}(M/\mathcal{R})$ at time t in the Galilean reference frame \mathcal{R} (the laboratory frame). From t to $t + dt$, the trajectory of the particle in \mathcal{R} is described by $dM = \mathbf{v}dt$. The space time interval between $M(t)$ and $M(t + dt)$ writes

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2 - c^2(dt)^2.$$

Since the space time interval is an invariant through a Lorentz transformation, it can also be calculated in the tangent Galilean frame \mathcal{R}_0 :

$$(ds)^2 = -c^2(dt_0)^2 \quad \text{by definition of the proper time.}$$

which defines the *proper time* for the particle as a measure of the motion of the particle along its space-time trajectory.

Using the invariance of the space-time interval, we can write

$$v^2(dt)^2 - c^2(dt)^2 = -c^2(dt_0)^2,$$

that is

$$(dt)^2 = \frac{c^2}{c^2 - v^2} (dt_0)^2 \Rightarrow dt = \frac{dt_0}{\sqrt{1 - v^2/c^2}}$$

We can then define a Lorentz factor of the particle, $\gamma(v) = \frac{1}{\sqrt{1 - v^2(t)/c^2}}$, using its instantaneous velocity $v(t)$ in order to find a relation between time in the laboratory and the proper time of the particle

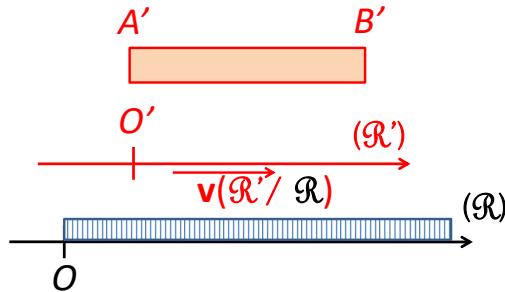
$$dt = \gamma dt_0.$$

The proper time is therefore the time measured by a clock in its own rest frame.

13.3.5 Length contraction

- PROPER LENGTH

The *proper length* of an object denotes its length in the Galilean reference frame where it is at rest.



Consider an object $A'B'$ at rest in the reference frame \mathcal{R}' , moving at velocity v with respect to \mathcal{R} . An observer in \mathcal{R}' measures with a ruler the length $\Delta x'$ of the the object $A'B'$. $\Delta x'$ is the proper length of the object.

An observer in the laboratory frame \mathcal{R} observes the coincidence of the object ends with fixed points of \mathcal{R} at the same instant t and measures the distance between these fixed points with a ruler at rest in \mathcal{R} . The observer in the laboratory reports a shorter length than the observer in the moving frame. He concludes that moving objects contract in the direction of motion.

- LENGTH CONTRACTION EQUATION

The length measured by the observer in \mathcal{R} and the proper length in \mathcal{R}' are linked by the Lorentz transformation:

$$x' = \gamma(x - vt),$$

that can be applied to the positions of the ends of the object, measured at the same time in \mathcal{R} , that is, $\Delta t = 0$.

$$\Delta x' = \gamma(\Delta x - v\Delta t).$$

We finally find

$$\Delta x' = \gamma \Delta x,$$

Since $\gamma \geq 1$, we have $\Delta x < \Delta x'$ which confirms that the observer in the laboratory reports a shorter length than the observer in the moving frame.

Length contraction is a kinematic effect. It is not the result of any compression force that would occur due to the motion of the object. Rather, it is again a manifestation of the non-absolute nature of time in a Universe without infinite velocity.

13.3.6 Relativistic velocity addition

Alice is onboard a rocket, flying at relativistic speed equal to $0.9c$, toward Bob who is waiting for her, sitting on a remote star (assumed at rest). She sends him a message using radio waves. What is the speed of the message, relative to Bob's star?

If Alice and Bob were living in a Newtonian world, the message would travel at $0.9c + c = 1.9c$, which is obviously wrong since c , the velocity of light, is the highest possible velocity for the radio-wave signal to travel. The composition law for velocities must be adapted to be compatible with the Lorentz transformation.

Consider the velocity $\mathbf{v}(M/\mathcal{R})$ of a material point M in the Galilean frame \mathcal{R} . By definition of a velocity, it is given by differentiating the position with respect to time, which yields the three components

$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt} \quad v_z = \frac{dz}{dt}.$$

Now consider a Galilean reference frame \mathcal{R}' , moving at velocity $\mathbf{u}(\mathcal{R}'/\mathcal{R})$ with respect to \mathcal{R} . Without loss of generality, we can assume that the axes of \mathcal{R} and \mathcal{R}' are chosen such that the x -axis is parallel to the x' -axis and to \mathbf{u} . The transverse axes are also parallel, that is, $Oy \parallel O'y'$ and $Oz \parallel O'z'$.

In the frame \mathcal{R}' , the velocity of the material point is $\mathbf{v}'(M/\mathcal{R}')$, defined by

$$v'_{x'} = \frac{dx'}{dt'}, \quad v'_{y'} = \frac{dy'}{dt'} \quad v'_{z'} = \frac{dz'}{dt'}.$$

Now we can use the Lorentz transformation to express the velocity components in \mathcal{R}' as functions of those in \mathcal{R} . For the component along the x -axis, we find

$$\begin{aligned} dx' &= \gamma(dx - u dt) \\ dt' &= \gamma \left(dt - \frac{udx}{c^2} \right) \end{aligned}$$

leading to

$$v'_{x'} = \frac{dx - u dt}{dt - \frac{udx}{c^2}} \rightarrow v'_{x'} = \frac{v_x - u}{1 - \frac{uv_x}{c^2}}$$

We see that the numerator is just what we would expect in a Newtonian world but the denominator brings a correction. Going back to the question on the velocity of Alice's message to Bob, we have $u = 0.9c$, $v'_{x'} = c$ and the message velocity is obtained as

$$v_x = \frac{v'_{x'} + u}{1 + \frac{uv'_{x'}}{c^2}} = \frac{c + 0.9c}{1 + \frac{0.9cc}{c^2}} = c$$

As it should, the radio wave carrying Alice's message travels at c in Bob's reference frame. For the velocity components in the transverse directions, we find

$$\begin{cases} dy' = dy, \\ dz' = dz \end{cases} \rightarrow \begin{cases} \frac{dy'}{dt'} = \frac{dy}{\gamma \left(dt - \frac{udx}{c^2} \right)} \\ \frac{dz'}{dt'} = \frac{dz}{\gamma \left(dt - \frac{udx}{c^2} \right)} \end{cases}$$

leading to

$$v'_{y'} = \frac{v_y}{\gamma \left(1 - \frac{uv_x}{c^2} \right)}, \quad v'_{z'} = \frac{v_z}{\gamma \left(1 - \frac{uv_x}{c^2} \right)}.$$

We see that the velocity components transverse to the direction of motion of the frame \mathcal{R}' are also modified. This is not the case in a Newtonian Universe where only the longitudinal velocity component is modified. Again, this is a kinematic effect due to the relative nature of time.

We also note that the equivalence principle is verified, that is, the composition rule for velocities known in Newtonian mechanics is retrieved for $u \ll c$:

$$v'_{x'} \rightarrow v_x - u, \quad v'_{y'} \rightarrow v_y, \quad v'_{z'} \rightarrow v_z.$$

For an electromagnetic signal (satisfying $v_x'^2 + v_y'^2 + v_z'^2 = c^2$), it is easy to verify that

$$v'_{x'}^2 + v'_{y'}^2 + v'_{z'}^2 = c^2.$$

Indeed, squaring the velocity components and replacing $1/\gamma^2$ by $(1 - u^2/c^2)$, we find

$$\begin{aligned} v'_{x'}^2 + v'_{y'}^2 + v'_{z'}^2 &= \frac{(v_x - u)^2}{\left(1 - \frac{uv_x}{c^2} \right)^2} + \frac{v_y^2}{\gamma^2 \left(1 - \frac{uv_x}{c^2} \right)^2} + \frac{v_z^2}{\gamma^2 \left(1 - \frac{uv_x}{c^2} \right)^2} \\ &= \frac{(v_x^2 - 2uv_x + u^2) + (v_y^2 + v_z^2)(1 - u^2/c^2)}{\left(1 - \frac{uv_x}{c^2} \right)^2} \\ &= \frac{v_x^2 - 2uv_x + u^2 + (c^2 - v_x^2) - u^2(1 - v_x^2/c^2)}{(1 - 2uv_x/c^2 + u^2v_x^2/c^4)} = c^2. \end{aligned}$$

• SIMULTANEITY AND LOCALIZATION

It results from the Lorentz transformation that space intervals and time intervals in \mathcal{R} and in \mathcal{R}' are linked by the relations

$$\begin{aligned}\Delta x' &= \gamma(\Delta x - u\Delta t) \\ \Delta t' &= \gamma(\Delta t - u\Delta x/c^2)\end{aligned}$$

where u denotes the velocity of \mathcal{R}' with respect to \mathcal{R} .

Thus, if two events are localized at the same position in \mathcal{R} ($\Delta x = 0$), an observer of \mathcal{R}' does not observe them at the same position as

$$\Delta x' = -\gamma u \Delta t \neq 0.$$

Similarly, if two events are simultaneous in \mathcal{R} ($\Delta t = 0$), an observer of \mathcal{R}' observes them different times as

$$\Delta t' = -\gamma u \Delta x/c^2 \neq 0.$$

13.3.7 Space-time interval between two events

We have seen that the space-time interval between two events is an invariant. This means that it is invariant to translations and rotations in space, invariant to translations in time, and invariant to a Lorentz transformation. Hence, the interval

$$(\Delta s)^2 = \underbrace{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}_{d^2} - (c^2 \Delta t)^2$$

takes the same value in all Galilean reference frames.

Since $(\Delta s)^2$ can take different signs, this leads to a classification of the intervals between two events, depending whether the spatial distance d is smaller or larger than the distance an electromagnetic signal can cover during the time interval Δt . We thus have three classes:

$$\begin{array}{lll} (\Delta s)^2 < 0 & \text{time-like separation} & d < c\Delta t \\ (\Delta s)^2 = 0 & \text{null separation} & d = c\Delta t \\ (\Delta s)^2 > 0 & \text{space-like separation} & d > c\Delta t \end{array}$$

which receives the following interpretation:

- Events with a null separation $d = c\Delta t$ can be connected by a signal traveling at the speed of light.
- Two events with a space-like separation are separated by a greater distance than the distance an electromagnetic signal can cover during Δt , that is $|c\Delta t/\Delta x| < 1$. It is then always possible to perform a Lorentz transformation to a Galilean frame \mathcal{R}' where the events are simultaneous ($\Delta t' = 0$) with

$$(\Delta s')^2 = (\Delta x')^2 - (c\Delta t')^2 = (\Delta x')^2.$$

This transformation is characterized by its normalized velocity $\beta = |c\Delta t/\Delta x|$.

- For events separated by a time-like interval, the distance in space is smaller than the distance an electromagnetic signal can cover during Δt . It is possible to perform a Lorentz transformation so that the events are localized at the same position ($\Delta x' = 0$) with

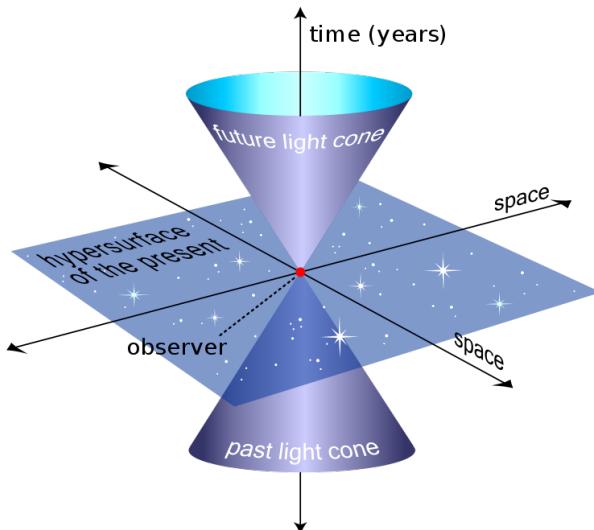
$$(\Delta s')^2 = (\Delta x')^2 - (c\Delta t')^2 = -c^2(\Delta t')^2.$$

This transformation is characterized by its normalized velocity $\beta = |\Delta x/c\Delta t|$.

This classification leads to the space-time Minkowski diagram shown in the figure below, which is a representation of the interval classes in the 4-dimensional space-time, using a single variable for two spatial variables. For instance, using $\rho^2 = y^2 + z^2$, we obtain two cones for the class of null separation, of equation $\rho^2 + x^2 = c^2t^2$, which is called the light cone. The light cone separates the class of space-like intervals (outside the light cone, in the direction of space axes) from the class of time-like intervals (inside the light cone, in the direction of time).

Consider an event (observer) occupying the origin of coordinates, $t = 0$, $x = 0$, $\rho = 0$. The plane that comprises spatial coordinates is the hyperplane of the present. Any event, say T , inside the light cone lies either inside the future light cone and will occur after the event at O in all Galilean frames, or inside the past light cone and occurred before the event at O in all Galilean frames. The time-like interval between the events O and T make it possible for a cause and effect relationship to exist between them, that is, for a signal of velocity smaller than the light velocity to travel between the two events.

In contrast, for a space-like interval, that is, an event S lying outside the light cone, a cause and effect relationship between O and S is not possible. The time interval between these events may have different signs for different observers. No signal propagating at a velocity smaller than the light velocity can travel between O and S . It is impossible to transform these events to the same point in space as would be required to compare their clocks, if we wanted to check a cause and effect relationship.



13.3.8 Four vectors

Einstein's postulate of relativity states that the laws of physics have the same form in every Galilean frame. This invariance can be formalized using a suited mathematical framework which was introduced by Minkowsky and is presented in this section.

• EUCLIDIAN SPACE

In the usual *Euclidian space*, vectors are defined by their three spatial components, for instance, $\mathbf{a} = (a_1, a_2, a_3)$, and $\mathbf{b} = (b_1, b_2, b_3)$. We will call them *three-vectors*. The scalar product between the three-vectors \mathbf{a} and \mathbf{b} is defined by

$$\mathbf{a} \cdot \mathbf{b} = a_k b_k \equiv \sum_{k=1}^3 a_k b_k,$$

where the repeated subscripts k means that a summation over k is performed. We will use this convention, called the Einstein convention, in the remaining of the relativity course.

The scalar product is invariant to translations and rotations of the system of coordinates. If \mathbf{a}' and \mathbf{b}' denotes the new components of \mathbf{a} and \mathbf{b} after a rotation or a translation, we have

$$\mathbf{a} \cdot \mathbf{b} = a_k b_k = a'_k b'_k = \mathbf{a}' \cdot \mathbf{b}'$$

The norm $\sqrt{\mathbf{a} \cdot \mathbf{a}}$ is also invariant to translations or rotations of the coordinate system in an Euclidian space, this property being a special case of the invariance of the scalar product.

• MINKOWSKI'S SPACE

For special relativity, Minkowski's space plays the role of the Euclidian space for Newtonian mechanics. The space-time framework is naturally described by four coordinates and thus by *four-vectors*

$$\mathbf{a} = (a_1, a_2, a_3, a_4).$$

A prototypical four-vector in special relativity is the space-time coordinate vector

$$\vec{r} = (x, y, z, ict),$$

where the fourth component is purely imaginary.

In Minkowsky's space, the scalar product between \mathbf{a} and $\mathbf{b} = (b_1, b_2, b_3, b_4)$ is defined in the same way as in an Euclidian space, just with one more component, and according to the usage, component subscripts are written with a greek letter, which is repeated to indicate a sum over the subscript values. The scalar product is invariant to translations, rotations and Lorentz transformations from a Galilean frame to another, that is,

$$\mathbf{a} \cdot \mathbf{b} = a_\mu b_\mu = a'_\mu b'_\mu = \mathbf{a}' \cdot \mathbf{b}'.$$

Still in analogy with the Euclidian space, the norm $\sqrt{\mathbf{a} \cdot \mathbf{a}}$ is an invariant as it is a particular case of an invariant scalar product. Now we see that applied to space-time coordinates, this yields the invariant space-time interval

$$(\Delta s)^2 = \Delta \vec{r} \cdot \Delta \vec{r} = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2,$$

as the imaginary i in the fourth component gives the minus sign in the calculation of the space-time interval as a dot product.

• LORENTZ TRANSFORMATION IN 4-VECTOR NOTATION

Four-vectors are particularly useful in writing the Lorentz transformation in a compact form of a matrix-vector product, using the convention of repeated subscripts for a sum. The Lorentz transformation between the Galilean frames \mathcal{R} and \mathcal{R}' , and its inverse indeed read

$$r'_\mu = L_{\mu\nu} r_\nu, \quad \text{and} \quad r_\mu = L_{\mu\nu}^{-1} r'_\nu,$$

where $L_{\mu\nu}$ and $L_{\mu\nu}^{-1}$ denotes the 4×4 Lorentz matrix and its inverse:

$$L = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \quad L^{-1} = \begin{bmatrix} \gamma & 0 & 0 & -i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ i\beta\gamma & 0 & 0 & \gamma \end{bmatrix}$$

It is assumed, as usual in this chapter, that the motion of \mathcal{R}' is along $(Ox') \parallel (Ox)$ at velocity βc with respect to \mathcal{R} .

The expression for the inverse of the Lorentz matrix shows that L is an orthogonal matrix: $L^T = L^{-1}$. The determinant of the Lorentz matrix is one: $|L| = |L^{-1}| = 1$.

Now, we can apply a Lorentz transformation more generally to an arbitrary four vector \mathbf{a} whose space-components are a_1, a_2, a_3 and whose time-component is a_4 . In fact, the very definition of a four-vector of Minkowsky's space is precisely that it is transformed exactly like the space-time coordinate \vec{r} :

$$a'_\mu = L_{\mu\nu} a_\nu$$

13.3.9 General Lorentz's transformation

So far, we considered the Lorentz transformation between two Galilean frames \mathcal{R} and \mathcal{R}' where coordinate axes in \mathcal{R} and \mathcal{R}' were chosen parallel and orientation of the Ox and $O'x'$ axes parallel to \mathbf{v} , the velocity of \mathcal{R}' with respect to \mathcal{R} . We might need to express the Lorentz transformation between two Galilean frames in a more general configuration, that is, in the case where \mathbf{v} is not aligned with any axis of the frames, or in the case where the Cartesian axes of the two frames are not aligned. In the latter case, it is possible to perform first a rotation of axes using an Euler angle transformation to align the axes. We then end up with aligned Cartesian axes between the two frames and with an arbitrary orientation of the vector quantity

$$\beta \equiv \mathbf{v}/c,$$

which represents the dimensionless velocity of \mathcal{R}' with respect to \mathcal{R} . Then, decompose any vector \mathbf{a} into its components \mathbf{a}_{\parallel} and \mathbf{a}_{\perp} that are parallel and perpendicular to β , respectively. The Lorentz transformation and its inverse takes simple expressions using this decomposition:

$$\begin{aligned} \mathbf{a}'_{\perp} &= \mathbf{a}_{\perp} & \mathbf{a}_{\perp} &= \mathbf{a}'_{\perp} \\ \mathbf{a}'_{\parallel} &= \gamma(\mathbf{a}_{\parallel} + i\beta a_4) & \mathbf{a}_{\parallel} &= \gamma(\mathbf{a}'_{\parallel} - i\beta a'_4) \\ a'_4 &= \gamma(a_4 - i\beta \cdot \mathbf{a}_{\parallel}) & a_4 &= \gamma(a'_4 + i\beta \cdot \mathbf{a}'_{\parallel}) \end{aligned}$$

Note that these expressions assume aligned axes between the two frames. The most general Lorentz transformation requires applying first the Euler angle transformation to align the axes. One can choose the order of the rotation of axes and Lorentz's transformation, but the β vector and the Euler angles must be adapted as they depend on the chosen order.

13.4 Relativistic dynamics

In this section, we will see other examples of four-vectors and we will apply the formalism of four-vectors to formulate the laws for the dynamics of relativistic particles, which extend Newton's laws of classical mechanics to the case of particle velocities approaching the velocity of light.

The laws must be covariant, that is invariant to a Lorentz transformation. They must also satisfy the equivalence principle, that is, Newton's laws must be retrieved for velocities that are small compared with the velocity of light.

13.4.1 Four-velocity and four-acceleration

• FOUR-VELOCITY

We start with another example of four-vector, the four-velocity, that will be useful in the context of dynamics of relativistic particles. In an Euclidian space, the three-velocity is obtained by differentiating in time the position three-vector. Since we wish to obtain a four-velocity \mathbf{U} that represents a relativistic version of the velocity of a particle, we will differentiate the four-position with respect to the proper time of the particle, that is, we take into account the Lorentz factor and write

$$\vec{U} = \underbrace{\frac{d\vec{r}}{dt_0}}_{\text{Proper time}} = \gamma(v) \frac{d}{dt}(x, y, z, ict) = \gamma(v) \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}, ic \right) = \gamma(v)(\mathbf{v}, ic) = (U_1, U_2, U_3, U_4)$$

This tells us that \mathbf{U} is a time-like four-vector, as the squared norm of \mathbf{U} is equal to a negative invariant quantity:

$$\vec{U} \cdot \vec{U} = (U_1^2 + U_2^2 + U_3^2 + U_4^2) = \gamma^2(v^2 - c^2) = -c^2.$$

It is therefore always possible to find a Galilean frame in which the three-velocity of the particle \mathbf{v} is instantaneously zero.

• FOUR-ACCELERATION

Differentiating the four-velocity with respect to the proper time, we obtain the four-acceleration

$$\mathcal{A} = \frac{d\mathbf{U}}{dt_0} = \gamma(v) \frac{d}{dt} \gamma(v)(\mathbf{v}, ic) = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4)$$

After differentiation, we find

$$\mathbf{A} = \underbrace{(\gamma^2 \mathbf{a} + \gamma^4 \mathbf{v}(\mathbf{v} \cdot \mathbf{a})/c^2)}_{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3}, \underbrace{i\gamma^4(\mathbf{v} \cdot \mathbf{a})/c}_{\mathcal{A}_4},$$

where $\mathbf{a} = d\mathbf{v}/dt$ denotes the acceleration three-vector.

13.4.2 Relativistic momentum and energy

• FOUR MOMENTUM AND ENERGY

A scalar quantity \mathcal{E} , the energy of the system, can be associated with any physical system. This quantity is a state function in the sense that its variation only depends on the initial and final states of the system.

If the system is isolated, its energy is constant. We are looking for an expression of the energy that is compatible with the relativity principle. Energy satisfies a conservation law, and a law for the conservation of momentum will be found as a consequence of the conservation of energy. We will build another four-vector, the four-momentum \mathbf{P} , and identify from its expression the relativistic energy and momentum that are compatible with the relativity principle. It is only after this identification that we will know that components of \mathbf{P} indeed correspond to a momentum or an energy. The four-momentum will play an important role in the dynamics of relativistic particles as it will allow us to extend Newton's law to the relativistic regime.

We define \mathbf{P} by multiplying the four-velocity by a mass m . We postulate:

The mass m is a Lorentz invariant quantity.

That is, invariant to a Lorentz transformation. The four-momentum then reads

$$\mathbf{P} = m\mathbf{U} = m\gamma(\mathbf{v}, ic).$$

Writing \mathbf{P} as

$$\mathbf{P} = (\mathbf{p}, P_4),$$

we recognize in the first three components the three-momentum

$$\mathbf{p} = m(U_1, U_2, U_3) \rightarrow \boxed{\mathbf{p} = \gamma m \mathbf{v}},$$

after performing a Taylor expansion for $v \ll c$:

$$\mathbf{p} = \gamma m \mathbf{v} = m \mathbf{v} \left[1 + \frac{1}{2} \frac{v^2}{c^2} + \frac{3}{8} \frac{v^4}{c^4} + \dots \right]$$

which shows that \mathbf{p} reduces to its Newtonian expression for small particle velocities with respect to the light velocity.

We write the fourth component as $P_4 = i\mathcal{E}/c$, where

$$\mathcal{E}/c = -imU_4 = \gamma mc \quad \rightarrow \quad \boxed{\mathcal{E} = \gamma mc^2}.$$

Hence, we can interpret the fourth component of the four-momentum vector as the ratio of an energy by the light velocity. Indeed, performing a Taylor expansion of \mathcal{E} for $v \ll c$, we find

$$\mathcal{E} = \frac{mc^2}{\sqrt{1-v^2/c^2}} = \underbrace{mc^2}_{\text{Rest energy}} + \underbrace{\frac{1}{2}mv^2}_{\text{Classical } E_{\text{kin}}} + \frac{3}{8}m\frac{v^4}{c^2} + \dots$$

where, by neglecting the higher-order terms of the expansion, we recognize the sum of a rest energy, when $v = 0$, and the classical kinetic energy $mv^2/2$. The four-momentum is thus a four-vector that generalizes to the relativistic regime the classical momentum and total energy of a particle.

The rest energy corresponds to the famous Einstein formula

$$\mathcal{E} = mc^2,$$

which tells us that the mass of a particle is equivalent to its energy.

The consequence of this equivalence is that the mass defect of a nucleus, defined as the difference between the mass of a nucleus and the sum of the masses of the nucleons within the nucleus, represents the energy binding the nucleus.

$$\Delta m = m_{\text{nucleus}} - \sum_k m_{k,\text{nucleons}} < 0$$

The negative sign of the mass defect means that the nucleus can liberate the corresponding amount of energy, $(\Delta m)c^2$ if the nucleons are separated. This is the basis of the nuclear energy technology by fission of heavy nuclei (Uranium).

A second consequence is that in relativistic dynamics, mass is not conservative as it can be exchanged with energy. Mass is however invariant to a Lorentz transformation.

• KINETIC ENERGY

The kinetic energy T for a relativistic particle can be defined by expressing the total energy $\mathcal{E} = \gamma mc^2$ as the sum of the rest energy and the relativistic kinetic energy. Hence,

$$T = \mathcal{E} - mc^2 = (\gamma - 1)mc^2$$

13.4.3 Relation between momentum and energy

The expressions for the relativistic momentum and energy allows us to calculate the squared norm of the four-momentum:

$$\begin{aligned} \mathbf{P} \cdot \mathbf{P} &= \mathbf{p}^2 - \frac{\mathcal{E}^2}{c^2} = (\gamma m\mathbf{v})^2 - \frac{(\gamma mc^2)^2}{c^2} \\ &= -m^2c^2. \end{aligned}$$

This is in keeping with the postulate that the mass is a Lorentz invariant.

From this invariance, we infer the relation between the relativistic momentum and energy:

$$\mathcal{E} = \sqrt{p^2 c^2 + m^2 c^4} = \gamma m c^2$$

Another useful relation is obtained from the ratio of the momentum to total energy, which yields the velocity of a relativistic particle

$$\mathbf{v} = \frac{c^2 \mathbf{p}}{\mathcal{E}}.$$

These relations apply the particular case of photons, which are massless particles. Indeed for $m = 0$, we find

$$\mathcal{E} = pc \quad \text{and} \quad \mathbf{v} = c\mathbf{p}/p,$$

which shows that photons travel at the velocity of light ($|\mathbf{v}| = c$).

13.4.4 Relativistic particle dynamics

Definition 13.4: Relativistic particle dynamics

The motion of a particle of velocity \mathbf{v} is characterized by the Lorentz factor $\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$. From the relativistic momentum $\mathbf{p} = \gamma m \mathbf{v}$, the equation of motion for a particle with charge q and mass m in an electromagnetic field is

$$\frac{d\mathbf{p}}{dt} = \frac{d}{dt}[\gamma m \mathbf{v}] = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

More generally,

$$\frac{d\mathbf{p}}{dt} = \mathbf{F},$$

where \mathbf{F} denotes the force applied to the particle. This law constitutes a postulate that extends Newton's law to the relativistic regime. Its validity was confirmed by experiments.

13.4.5 Conservation of power

The rate of change of the kinetic energy of a particle reads

$$\frac{dT}{dt} = \frac{d}{dt}[\mathcal{E} - mc^2] = \frac{d\mathcal{E}}{dt}.$$

From the relation between \mathcal{E} and \mathbf{p} ,

$$\mathcal{E}_2 = p^2 c^2 + m^2 c^4,$$

we obtain by differentiation

$$\mathcal{E} \frac{d\mathcal{E}}{dt} = c^2 \mathbf{p} \cdot \frac{d\mathbf{p}}{dt}.$$

Replacing $c^2 \mathbf{p}$ by $\mathcal{E} \mathbf{v}$

$$\frac{d\mathcal{E}}{dt} = \mathbf{v} \cdot \frac{d\mathbf{p}}{dt} = \mathbf{v} \cdot \mathbf{F},$$

we retrieve the kinetic energy theorem: the time rate of change of the kinetic energy of a particle moving at velocity \mathbf{v} is equal to the time rate of change of the work done on the particle by the force \mathbf{F} , that is, the power of force \mathbf{F} .

$$\frac{dT}{dt} = \mathbf{F} \cdot \mathbf{v}.$$

This law for conservation of power is the same as in classical mechanics provide the kinetic energy is replaced by its relativistic counterpart.

13.5 Summary and essential formulas

- The Lorentz transformation is the transformation law between the Galilean frames \mathcal{R} and \mathcal{R}' , moving at velocity \mathbf{v} with respect to \mathcal{R} . Using Cartesian axes with $(Ox) \parallel (O'x') \parallel \mathbf{v}$, $(Oy) \parallel (O'y')$ and $(Oz) \parallel (O'z')$, the Lorentz transformation reads

$$\begin{bmatrix} x' \\ y' \\ z' \\ ict' \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ ict \end{bmatrix}$$

where $\beta = \frac{v}{c} \leq 1$ and $\gamma = \frac{1}{\sqrt{1 - \beta^2}} \geq 1$.

- A four-vector is a vector of the four-dimensional Minkowsky space whose coordinates transform according to the Lorentz transformation between two Galilean frames.
- The four-momentum $\mathbf{P} = (\mathbf{p}, i\mathcal{E}/c)$ transforms according to Lorentz's transformation.
- For a particle of mass m and velocity v , the relativistic momentum and energy read

$$\begin{aligned} \text{Momentum} \quad \mathbf{p} &= \gamma m \mathbf{v} \\ \text{Energy} \quad \mathcal{E} &= \gamma mc^2 = \sqrt{p^2 c^2 + m^2 c^4} \end{aligned}$$

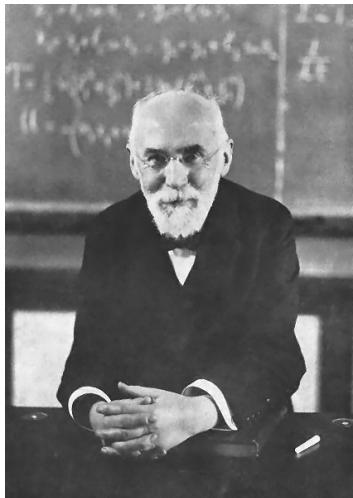
where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the Lorentz factor associated with the particle.

- The dynamics of a relativistic particle of mass m and charge q that is subject to an electromagnetic field \mathbf{E} , \mathbf{B} , satisfies the equation of motion

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

Chapter 14

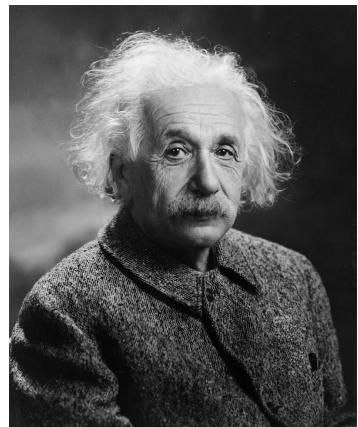
Covariant electrodynamics



Lorentz



Minkowski



Einstein

14.1 Introduction

In the previous chapter, we have seen how the results of Michelson and Morley's experiments have led the scientists of the XIXth century to abandon the hypothesis of an absolute reference frame (Aether) for the validity of Maxwell's equations, in favor of a reconstruction of the concepts of space and time governing the laws of physics. We presented Minkowsky's space and the Lorentz transformation between two Galilean frames, which are the two fundamental building blocks that led Einstein to the theory of special relativity.

The theory of special relativity is based on the relativity principle recalled in the next section, which allowed Einstein to reconstruct the kinematics at stake in our world and to generalize Newton's classical mechanics to the case of particles approaching the velocity of light, that is, reaching relativistic velocities. Newton's mechanics can be seen as a limit case of special relativity for particles moving at velocity $v \ll c$. This is the equivalence principle.

The purpose of this chapter is to show how Einstein's relativity principle applies to Maxwell's equations. In other words, the goal is to show that Maxwell's equations take the same form in every Galilean frame. This is not obvious a priori. As will be seen in section 14.5, it is possible to perform a cumbersome direct application of the Lorentz transformation between two Galilean frames. This will lead to the expected result. However, a more elegant approach, called covariant formulation of the laws of electrodynamics, will lead to an expression of Maxwell's equations that is manifestly invariant with respect to a change of Galilean frame. The presentation of this formulation will conclude this course.

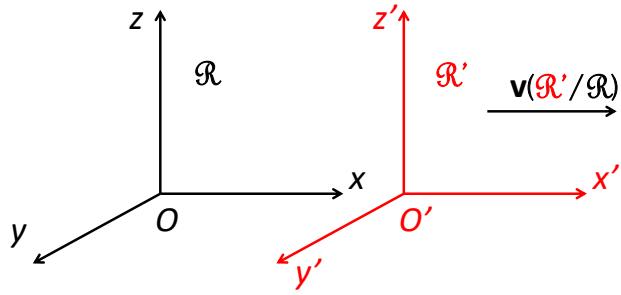
14.1.1 Einstein special relativity principle (1905)

We start with a brief reminder of Einstein's relativity principle:

- Relativity principle:
 - The fundamental laws of physics are covariant with respect to a change of Galilean reference frame, that is, the laws of physics take the same form in all Galilean frames. This generalizes Galileo's relativity principle so as to include Maxwell's equations.
 - The velocity of light in vacuum is the same in all Galilean frames. Light propagates in vacuum isotropically at velocity $c = 1/\sqrt{\epsilon_0\mu_0}$. This velocity is the velocity limit that a particle cannot exceed and the limit for the propagation of a signal carrying information or energy.

14.1.2 Lorentz's transformation

In this section, we recall the expressions for the Lorentz transformation that will be extensively used in the rest of this chapter.



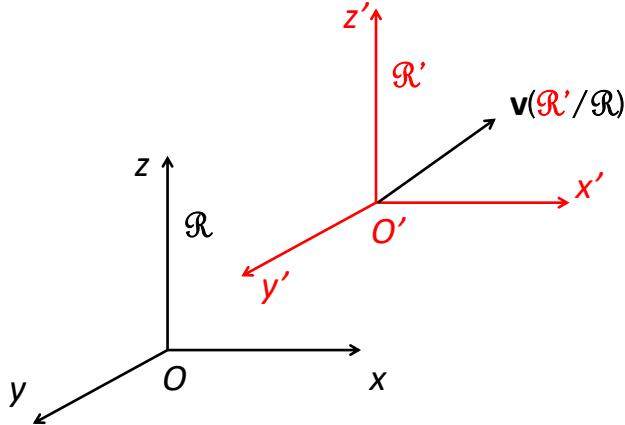
For a Galilean frame \mathcal{R}' moving at velocity \mathbf{v} with respect to the Galilean frame \mathcal{R} , with aligned Cartesian axes such that $(Ox) \parallel (O'x') \parallel \mathbf{v}$, $(Oy) \parallel (O'y')$ and $(Oz) \parallel (O'z')$, the Lorentz transformation and its inverse read

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma \left(t - \frac{vx}{c^2} \right) \end{aligned}$$

$$\begin{aligned} x &= \gamma(x' + vt) \\ y &= y' \\ z &= z' \\ t &= \gamma \left(t' + \frac{vx'}{c^2} \right) \end{aligned}$$

where γ denotes the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$



The velocity \mathbf{v} of \mathcal{R} with respect to \mathcal{R}' may not be aligned with the x - and x' -axes. In this case, and if the Cartesian axes of \mathcal{R} and \mathcal{R}' are still aligned¹, a more general expression of the Lorentz transformation reads

$$\begin{cases} \mathbf{r}'_\perp = \mathbf{r}_\perp \\ \mathbf{r}'_\parallel = \gamma(\mathbf{r}_\parallel - \beta ct) \\ ct' = \gamma(ct - \beta \cdot \mathbf{r}_\parallel) \end{cases} \quad \begin{cases} \mathbf{r}_\perp = \mathbf{r}'_\perp \\ \mathbf{r}_\parallel = \gamma(\mathbf{r}'_\parallel + \beta ct') \\ ct = \gamma(ct' + \beta \cdot \mathbf{r}'_\parallel) \end{cases} \quad (14.1)$$

¹Rotations can be applied to align axes if needed.

where the space-coordinates are projected on the parallel or perpendicular directions to \mathbf{v} , and the dimensionless parameter $\beta = v/c$, or its vector counterpart

$$\boldsymbol{\beta} = \frac{\mathbf{v}}{c}$$

is called the *boost*. The frame \mathcal{R}' is often called the *boosted frame*. As previously, γ denotes the Lorentz parameter

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

14.1.3 Lorentz's transformation in four-vector notation

We have seen in the last chapter that convenient mathematical tools for algebra in Minkowski's space are four-vectors, which are vectors with three space-like coordinates and one time-like, purely imaginary, coordinate that transforms like space-time under a Lorentz transformation.

We will use the notation with arrows for four-vectors, and keep the boldface notation for three-vectors or the three space-like coordinates of a four-vector. For instance, the space-time vector reads

$$\vec{r} = (x, y, z, ict) = (\mathbf{r}, ict)$$

The scalar-product of two four-vectors and the norm of a four vector are invariant under a Lorentz transformation.

In particular, the space time interval is an invariant:

$$(\Delta s)^2 = \Delta \vec{r} \cdot \Delta \vec{r} = \Delta \vec{r}' \cdot \Delta \vec{r}' = (\Delta s')^2.$$

Using the four-vector notation, the Lorentz transformation or its inverse can be written

$$r'_\mu = L_{\mu\nu} r_\nu, \quad r_\mu = L_{\mu\nu}^{-1} r'_\nu, \quad \mu = 1, \dots, 4$$

where the repeated ν index means an implicit sum over this index (Einstein's convention). In the above equation, this simply means that the four-vector \mathbf{r}' is the product of the matrix L by the four-vector \mathbf{r} .

If the Cartesian axes of \mathcal{R} and \mathcal{R}' are aligned and the x - and x' -axes are parallel to $\boldsymbol{\beta}$, the Lorentz matrix L reads

$$L = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \quad L^{-1} = \begin{bmatrix} \gamma & 0 & 0 & -i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ i\beta\gamma & 0 & 0 & \gamma \end{bmatrix}$$

The Lorentz matrix is clearly an orthogonal matrix (its inverse is equal to its transpose and its determinant is equal to 1).

A four-vector is therefore a vector \vec{a} which transforms like \vec{r} under a Lorentz transformation, that is

$$a'_\mu(\vec{r}') = L_{\mu\nu} a_\nu(\vec{r}),$$

or equivalently, if the boost is not aligned with any of the basis vector used to express the four-vector components,

$$\beta = \frac{\mathbf{v}}{c}, \quad \begin{cases} \mathbf{a}'_{\parallel} = \gamma(\mathbf{a}_{\parallel} + i\beta a_4) \\ \mathbf{a}'_{\perp} = \mathbf{a}_{\perp} \\ a'_4 = \gamma(a_4 - i\beta \cdot \mathbf{a}_{\parallel}) \end{cases} \quad \begin{cases} \mathbf{a}_{\parallel} = \gamma(\mathbf{a}'_{\parallel} - i\beta a'_4) \\ \mathbf{a}_{\perp} = \mathbf{a}'_{\perp} \\ a_4 = \gamma(a'_4 + i\beta \cdot \mathbf{a}'_{\parallel}) \end{cases} \quad (14.2)$$

14.2 Electromagnetic quantities

In this section we will revisit the electromagnetic quantities such as the fields, potentials, distribution of charge or current density that allowed us to formulate Maxwell's equations. Our objective is to discover how these quantities are transformed via the Lorentz transformation.

Since Maxwell's equations are expressed using vector calculus operators such as the gradient, the divergence, the curl, etc, our first task is to perform a Lorentz transformation for derivatives with respect to space and time.

Consider the Galilean reference frame \mathcal{R} and another Galilean frame \mathcal{R}' moving at velocity \mathbf{v} with respect to \mathcal{R} . Assume the Cartesian axes of \mathcal{R} and \mathcal{R}' are aligned but the velocity \mathbf{v} is arbitrary. We can define the boost β :

$$\beta = \frac{\mathbf{v}}{c}.$$

Writing $\nabla \equiv \frac{\partial}{\partial \mathbf{r}}$, we use the chain rule and obtain the following derivatives

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}'_{\parallel}} &= \underbrace{\frac{\partial \mathbf{r}'_{\parallel}}{\partial \mathbf{r}'_{\parallel}}}_{\gamma} \cdot \frac{\partial}{\partial \mathbf{r}_{\parallel}} + \underbrace{\frac{\partial \mathbf{r}'_{\perp}}{\partial \mathbf{r}'_{\parallel}}}_{\mathbf{0}} \cdot \frac{\partial}{\partial \mathbf{r}_{\perp}} + \underbrace{\frac{\partial t}{\partial \mathbf{r}'_{\parallel}}}_{\gamma\beta/c} \frac{\partial}{\partial t} = \gamma \left(\frac{\partial}{\partial \mathbf{r}_{\parallel}} + \frac{\beta}{c} \frac{\partial}{\partial t} \right) \\ \frac{\partial}{\partial \mathbf{r}'_{\perp}} &= \frac{\partial}{\partial \mathbf{r}_{\perp}} \\ \frac{\partial}{\partial t'} &= \underbrace{\frac{\partial \mathbf{r}'_{\parallel}}{\partial t'}}_{\gamma c \beta} \cdot \frac{\partial}{\partial \mathbf{r}_{\parallel}} + \underbrace{\frac{\partial t}{\partial t'}}_{\gamma} \frac{\partial}{\partial t} = \gamma \left(c\beta \cdot \frac{\partial}{\partial \mathbf{r}_{\parallel}} + \frac{\partial}{\partial t} \right) \end{aligned}$$

We then use the expression of Lorentz's transformation (14.1) to recognize the Lorentz factor γ and the β vector, leading to

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}'_{\parallel}} &= \gamma \left(\frac{\partial}{\partial \mathbf{r}_{\parallel}} + \frac{\beta}{c} \frac{\partial}{\partial t} \right), \\ \frac{\partial}{\partial \mathbf{r}'_{\perp}} &= \frac{\partial}{\partial \mathbf{r}_{\perp}}, \\ \frac{\partial}{\partial t'} &= \gamma \left(c\beta \cdot \frac{\partial}{\partial \mathbf{r}_{\parallel}} + \frac{\partial}{\partial t} \right). \end{aligned}$$

Identifying the parallel and perpendicular components of the gradient operator

$$\nabla \equiv \frac{\partial}{\partial \mathbf{r}}, \quad \nabla_{\parallel} \equiv \frac{\partial}{\partial \mathbf{r}_{\parallel}}, \quad \nabla_{\perp} \equiv \frac{\partial}{\partial \mathbf{r}_{\perp}}, \quad \nabla'_{\parallel} \equiv \frac{\partial}{\partial \mathbf{r}'_{\parallel}}, \quad \nabla'_{\perp} \equiv \frac{\partial}{\partial \mathbf{r}'_{\perp}},$$

the transformation equations for partial derivatives take the form of a Lorentz transformation for the four gradient $\vec{\nabla}$:

$$\boxed{\begin{aligned}\nabla'_{\parallel} &= \gamma \left(\nabla_{\parallel} + i\beta \frac{\partial}{\partial(ict)} \right) \\ \nabla'_{\perp} &= \nabla_{\perp} \\ \frac{\partial}{\partial(ict')} &= \gamma \left(\frac{\partial}{\partial(ict)} - i\beta \nabla_{\parallel} \right)\end{aligned}}$$

$$\boxed{\vec{\nabla} = \left(\nabla, \frac{\partial}{\partial(ict)} \right).}$$

The norm of a four-vector is an invariant, thus the norm or the four-gradient is an covariant operator, invariant to a Lorentz transformation:

$$\vec{\nabla} \cdot \vec{\nabla} = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} = \nabla'^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t'^2}.$$

We see that the wave propagation operator (the d'Alembertian) is invariant to Lorentz transformations, that is, waves propagate at the speed of light in all Galilean frames, and we are now ready to check this invariance for propagation equations after expressing the transformation laws for the potentials and fields.

14.3 Continuity equation

Einstein's relativity principle states that the laws of electrodynamics are valid in every Galilean frame. This must be the case for the conservation of charge that we have written in the form of the continuity equation

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \quad (14.3)$$

The goal of this section is to show that Eq. (14.3) is indeed invariant to a Lorentz transformation. This equation can be rewritten using the four-gradient defined in the previous section and the vector²

$$\vec{j} = (\mathbf{j}, ic\rho) \quad (14.4)$$

The continuity equation then reads

$$\vec{\nabla} \cdot \vec{j} = 0, \quad (14.5)$$

where the operator $\vec{\nabla} \cdot$ denotes the four-divergence.

Now reading the resulting equation as the dot product of the four-gradient, which we have shown to be a four-vector, with \vec{j} defined in Eq. (14.4), we conclude that Eq. (14.5) is invariant to a Lorentz transformation only if the four-current $\vec{j} = (\mathbf{j}, ic\rho)$ is indeed a four-vector, that is, if it is invariant to a Lorentz transformation.

In order to check that \vec{j} is a four-vector, consider a charge element $dq' = \rho' dx' dy' dz'$, at rest in the Galilean frame \mathcal{R}' . In the frame \mathcal{R} , this charge element is expressed as $dq = \rho dx dy dz$, where the charge densities ρ and ρ' correspond to the densities viewed from their respective frame. If we assume that the direction of motion of \mathcal{R}' with respect to \mathcal{R} is along the x -axis (parallel to the x' -axis), there

²This is a definition. We do not know yet if it is a four-vector.

is length contraction and therefore, $dx = dx'/\gamma$. The electric charge is an invariant, which means $dq = dq'$. Thus, the charge densities in \mathcal{R} and in \mathcal{R}' must satisfy

$$dq = \rho dx dy dz = \rho \frac{dx'}{\gamma} dy' dz' = \frac{\rho}{\gamma} dx' dy' dz' = \rho' dx' dy' dz' = dq',$$

which implies density dilation

$$\rho = \gamma \rho'.$$

Finally, to be a four-vector, the four-current must transform according to the Lorentz transformation

$$\begin{aligned}\mathbf{j}'_{\parallel} &= \gamma(\mathbf{j}_{\parallel} - \rho \mathbf{v}), \\ \mathbf{j}'_{\perp} &= \mathbf{j}_{\perp}, \\ \rho' &= \gamma \left(\rho - \frac{jv}{c^2} \right).\end{aligned}$$

From the definition of the current $\mathbf{j} = \rho \mathbf{v}$, we can check that the equations for the parallel and perpendicular components are satisfied ($\mathbf{j}'_{\perp} = \mathbf{j}_{\perp} = 0$ and $\mathbf{j}'_{\parallel} = 0$ since the charge is at rest in \mathcal{R}'). For the last equation

$$\rho' = \gamma(1 - \beta^2)\rho = \frac{\rho}{\gamma},$$

as found above.

In conclusion \vec{j} is a four-vector and the continuity equation is frame invariant and reads

$$\vec{\nabla} \cdot \vec{j} = 0 \quad \text{with} \quad \vec{j} = (\mathbf{j}, ic\rho).$$

In any other Galilean frame \mathcal{R}' , the continuity equation reads

$$\vec{\nabla}' \cdot \vec{j}' = 0 \quad \text{with} \quad \vec{j}' = (\mathbf{j}', ic\rho').$$

14.4 Lorenz gauge potentials

In this section, we will show that the Lorenz gauge condition

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial V}{\partial t} = 0$$

is invariant to a Lorentz transformation³.

We first note that if we define the four-potential as

$$\vec{A} = \left(\mathbf{A}, i \frac{V}{c} \right),$$

³Incidentally, this section makes it clear that Ludvig Lorenz, a Dane physicist who gave his name to the gauge, is not the same physicist as Hendrik Lorentz (from the Netherlands), who gave his name to the force and to the transformation.

then the Lorenz gauge can be rewritten in the form of a four-divergence

$$\vec{\nabla} \cdot \vec{A} = 0,$$

or in other words, a dot product between the four-gradient and the four-potential \vec{A} . If we show that \vec{A} is indeed a four-vector, that is, if \vec{A} is transformed according to Eq. (14.2) under a Lorentz transformation between two Galilean frames, then the Lorenz gauge will be indeed invariant to a Lorentz transformation.

To check that \vec{A} is a four-vector, we recall that the potentials in the Lorenz gauge satisfy the inhomogeneous d'Alembert wave equations

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] V = -\frac{\rho}{\epsilon_0} \quad \text{and} \quad \left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \mathbf{A} = -\mu_0 \mathbf{j}$$

Thus, it is a straightforward step to conclude that the four-potential satisfies the inhomogeneous d'Alembert wave equation where the four-current is the source term:

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \left(\mathbf{A}, i \frac{V}{c} \right) = -\mu_0 (\mathbf{j}, ic\rho). \quad (14.6)$$

Now the operator $\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right]$, often called the d'Alembertian, is a Lorentz invariant operator: A Lorentz transformation from \mathcal{R} to \mathcal{R}' leads to

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \rightarrow \left[\nabla'^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} \right]$$

The four-current \vec{j} is a four-vector, which implies that the four-potential \vec{A} must be a four-vector as well, otherwise the transformation properties of left-hand-side and right-hand-side of Eq. (14.6) will differ.

In conclusion, the Lorenz gauge condition takes the same form in all Galilean frames.

14.5 Field transformation laws

The objective of this section is to express the electric and magnetic fields $\mathbf{E}'(\mathbf{r}', t')$ and $\mathbf{B}'(\mathbf{r}', t')$ viewed by an observer of a Galilean frame \mathcal{R}' moving at velocity $\mathbf{v} = c\beta$ with respect to \mathcal{R} , as a function of the fields $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ in \mathcal{R} .

For this task, we will exploit the definition of the fields

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t},$$

in conjunction with the transformation rules for the four-gradient $\vec{\nabla}$ and the four-potential \vec{A} .

Starting with the magnetic field, we perform a decomposition into the parallel and perpendicular components to \mathbf{v} :

$$\begin{aligned} \mathbf{B}' &= \nabla' \times \mathbf{A}' = (\nabla'_{||} + \nabla'_{\perp}) \times (\mathbf{A}'_{||} + \mathbf{A}'_{\perp}) \\ &= \nabla'_{||} \times \mathbf{A}'_{||} + \nabla'_{\perp} \times \mathbf{A}'_{\perp} + \nabla'_{||} \times \mathbf{A}'_{\perp} + \nabla'_{\perp} \times \mathbf{A}'_{||}. \end{aligned}$$

The first term vanishes $\nabla'_{\parallel} \times \mathbf{A}'_{\parallel} = \mathbf{0}$ and the remaining terms allow us to identify the parallel and perpendicular components of the field

$$\begin{aligned}\mathbf{B}'_{\parallel} &= (\nabla' \times \mathbf{A}')_{\parallel} = \nabla'_{\perp} \times \mathbf{A}'_{\perp}, \\ \mathbf{B}'_{\perp} &= (\nabla' \times \mathbf{A}')_{\perp} = \nabla'_{\parallel} \times \mathbf{A}'_{\perp} + \nabla'_{\perp} \times \mathbf{A}'_{\parallel}\end{aligned}$$

Both three-vectors ∇'_{\perp} and \mathbf{A}'_{\perp} are perpendicular components of the space-part of a four-vector, hence, they are invariant under a Lorentz transformation and so, the first equation tells us that the parallel component of the magnetic field is invariant:

$$\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}$$

For the perpendicular components of the field, we substitute the prime quantities using the Lorentz transformation:

$$\begin{aligned}\mathbf{B}'_{\perp} &= \gamma \left(\nabla_{\parallel} + \frac{\mathbf{v}}{c^2} \frac{\partial}{\partial t} \times \mathbf{A}_{\perp} + \nabla_{\perp} \right) \times \gamma \left(\mathbf{A}_{\parallel} - \frac{\mathbf{v}}{c^2} V \right) \\ &= \gamma \left(\nabla_{\parallel} \times \mathbf{A}_{\perp} + \nabla_{\perp} \times \mathbf{A}_{\parallel} \right) + \frac{\gamma}{c^2} \left(\mathbf{v} \times \frac{\partial \mathbf{A}_{\perp}}{\partial t} - \nabla_{\perp} \times (\mathbf{v} V) \right).\end{aligned}$$

We recognize \mathbf{B}_{\perp} in the first two terms of the right-hand-side, i.e., the perpendicular component of $\nabla \times \mathbf{A}$ if we perform a similar decomposition as for $\nabla' \times \mathbf{A}'$, and the fact that \mathbf{v} is a constant vector leads to

$$\mathbf{B}'_{\perp} = \gamma \mathbf{B}_{\perp} - \frac{\gamma \mathbf{v}}{c^2} \times \left(-\frac{\partial \mathbf{A}_{\perp}}{\partial t} - \nabla_{\perp} V \right)$$

Finally, from the definition of \mathbf{E} as a function of the potentials, we can write

$$\mathbf{B}'_{\perp} = \gamma \left(\mathbf{B}_{\perp} - \frac{\mathbf{v}}{c^2} \times \mathbf{E}_{\perp} \right) = \gamma \left(\mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E} \right)_{\perp},$$

or equivalently,

$$c \mathbf{B}'_{\perp} = \gamma (c \mathbf{B} - \beta \times \mathbf{E})_{\perp}.$$

For the electric field, we start from the parallel component

$$\mathbf{E}'_{\parallel} = -\nabla'_{\parallel} V' - \frac{\partial \mathbf{A}'_{\parallel}}{\partial t'}$$

and we use the inverse Lorentz transformation to substitute all prime quantities:

$$\mathbf{E}'_{\parallel} = -\gamma \left(\nabla_{\parallel} + \frac{\mathbf{v}}{c^2} \frac{\partial}{\partial t} \right) \gamma (V - \mathbf{v} \cdot \mathbf{A}) - \gamma \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \gamma \left(\mathbf{A}_{\parallel} - \frac{\mathbf{v}}{c^2} V \right).$$

Expanding the right-hand-side, we find

$$\mathbf{E}'_{\parallel} = \gamma^2 \left(-\nabla_{\parallel} V + \cancel{\nabla_{\parallel}(\mathbf{v} \cdot \mathbf{A})} - \frac{\mathbf{v}}{c^2} \cancel{\frac{\partial V}{\partial t}} + \frac{\mathbf{v}}{c^2} \frac{\partial(\mathbf{v} \cdot \mathbf{A})}{\partial t} - \frac{\partial \mathbf{A}_{\parallel}}{\partial t} + \cancel{\frac{\mathbf{v}}{c^2} \frac{\partial V}{\partial t}} - \cancel{(\mathbf{v} \cdot \nabla) \mathbf{A}_{\parallel}} + \frac{\mathbf{v}}{c^2} (\mathbf{v} \cdot \nabla) V \right)$$

Four terms cancel, including $\nabla_{\parallel}(\mathbf{v} \cdot \mathbf{A})$ and $-(\mathbf{v} \cdot \nabla) \mathbf{A}_{\parallel}$, since \mathbf{v} is a constant vector. The remaining terms can then be rewritten as

$$\mathbf{E}'_{\parallel} = -\gamma^2 \left(1 - \frac{v^2}{c^2} \right) \left(\nabla_{\parallel} V + \frac{\partial \mathbf{A}_{\parallel}}{\partial t} \right) = \left(\nabla_{\parallel} V + \frac{\partial \mathbf{A}_{\parallel}}{\partial t} \right) = \mathbf{E}_{\parallel}.$$

Finally, for the perpendicular component of the electric field, we have

$$\mathbf{E}'_{\perp} = -\nabla'_{\perp} V' - \frac{\partial \mathbf{A}'_{\perp}}{\partial t'}.$$

Again we substitute the prime quantities using the Lorentz transformation:

$$\begin{aligned}\mathbf{E}'_{\perp} &= -\nabla_{\perp} \gamma(V - \mathbf{v} \cdot \mathbf{A}) - \gamma \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \mathbf{A}_{\perp} \\ &= -\gamma \left(\nabla_{\perp} V + \frac{\partial \mathbf{A}_{\perp}}{\partial t} \right) + \gamma \nabla_{\perp} (\mathbf{v} \cdot \mathbf{A}) - \gamma (\mathbf{v} \cdot \nabla) \mathbf{A}_{\perp}\end{aligned}$$

The first two terms on the right-hand-side yield \mathbf{E}_{\perp} and we can use $\nabla = \nabla_{\parallel} + \nabla_{\perp}$ in the third term:

$$\mathbf{E}'_{\perp} = \gamma \mathbf{E}_{\perp} + \gamma \nabla (\mathbf{v} \cdot \mathbf{A}) - \gamma \nabla_{\parallel} (\mathbf{v} \cdot \mathbf{A}) - \gamma (\mathbf{v} \cdot \nabla) \mathbf{A}_{\perp}.$$

For two three-vectors \mathbf{a} and \mathbf{b} , we have the identity

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}).$$

Applying it to $\mathbf{a} = \mathbf{v}$ and $\mathbf{b} = \mathbf{A}$ and the fact that \mathbf{v} is a constant vector reduces the right-hand side to two terms

$$\nabla(\mathbf{v} \cdot \mathbf{A}) = (\mathbf{v} \cdot \nabla)\mathbf{A} + \mathbf{v} \times (\nabla \times \mathbf{A})$$

We know that $\nabla \times \mathbf{A} = \mathbf{B} = \mathbf{B}_{\parallel} + \mathbf{B}_{\perp}$. Taking the cross product with \mathbf{v} yields

$$\mathbf{v} \times (\nabla \times \mathbf{A}) = \mathbf{v} \times \mathbf{B}_{\perp}$$

We then replace $\nabla(\mathbf{v} \cdot \mathbf{A})$ into the expression for \mathbf{E}'_{\perp} :

$$\mathbf{E}'_{\perp} = \gamma \left(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}_{\perp} + (\mathbf{v} \cdot \nabla)\mathbf{A} - \nabla_{\parallel} (\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla) \mathbf{A}_{\perp} \right).$$

The last three terms in the latter equation cancel out since

$$\begin{aligned}(\mathbf{v} \cdot \nabla)\mathbf{A} &= (\mathbf{v} \cdot \nabla)\mathbf{A}_{\parallel} + (\mathbf{v} \cdot \nabla)\mathbf{A}_{\perp} \\ &= (\mathbf{v} \cdot \nabla_{\parallel})\mathbf{A}_{\parallel} + (\mathbf{v} \cdot \nabla)\mathbf{A}_{\perp} \\ &= \nabla_{\parallel}(\mathbf{v} \cdot \mathbf{A}_{\parallel}) + (\mathbf{v} \cdot \nabla)\mathbf{A}_{\perp} \\ &= \nabla_{\parallel}(\mathbf{v} \cdot \mathbf{A}) + (\mathbf{v} \cdot \nabla)\mathbf{A}_{\perp}.\end{aligned}$$

Hence, we see that the transformation for the perpendicular component of the field reads

$$\mathbf{E}'_{\perp} = \gamma (\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}_{\perp}),$$

or equivalently

$$\mathbf{E}'_{\perp} = \gamma (\mathbf{E}_{\perp} + \beta \times c \mathbf{B}_{\perp}).$$

Finally, the transformation law for the electromagnetic field and its inverse read

$$\begin{cases} \mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel} \\ \mathbf{E}'_{\perp} = \gamma(\mathbf{E} + \beta \times c \mathbf{B})_{\perp} \\ \mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel} \\ c \mathbf{B}'_{\perp} = \gamma(c \mathbf{B} - \beta \times \mathbf{E})_{\perp} \end{cases} \quad \begin{cases} \mathbf{E}_{\parallel} = \mathbf{E}'_{\parallel} \\ \mathbf{E}_{\perp} = \gamma(\mathbf{E}' - \beta \times c \mathbf{B}')_{\perp} \\ \mathbf{B}_{\parallel} = \mathbf{B}'_{\parallel} \\ c \mathbf{B}_{\perp} = \gamma(c \mathbf{B}' + \beta \times \mathbf{E}')_{\perp} \end{cases} \quad (14.7)$$

This transformation law does not take the form of the transformation law established for a four-vector. We note from these expressions that the electric or magnetic nature of a field are intrinsically observer dependent. In \mathcal{R}' , the electric field is a combination of the electric and magnetic fields in \mathcal{R} and so does the magnetic field in \mathcal{R}' . Observers in different Galilean frames will reach different conclusion regarding the electric or magnetic origin of an electromagnetic phenomenon. Nevertheless Equations (14.7) constitute the transformation rule that we need to apply, in conjunction with the Lorentz transformation for the four-gradient, in order to conclude on the invariance of Maxwell's equations under a Lorentz transformation.

14.6 Plane waves

Since electromagnetic waves propagate at the velocity of light in vacuum, it is instructive to perform a Lorentz transformation to a plane wave and check the transformation rules for their characteristics. Consider a monochromatic plane wave propagating in vacuum at velocity $c = \omega/k$ in the Galilean reference frame \mathcal{R} . The electromagnetic fields in \mathcal{R} read

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)], \quad c\mathbf{B} = \frac{\mathbf{k}}{k} \times \mathbf{E}.$$

Since the electric and magnetic fields in \mathcal{R} are solutions to d'Alembert's wave equation, and the d'Alembertian is invariant to a Lorentz transformation, a plane wave has exactly the same form if it is observed from another Galilean frame \mathcal{R}' moving at velocity \mathbf{v} with respect to \mathcal{R} . Hence, we can write

$$\mathbf{E}'(\mathbf{r}', t') = \mathbf{E}'_0 \exp[i(\mathbf{k}' \cdot \mathbf{r}' - \omega' t')], \quad c\mathbf{B}' = \frac{\mathbf{k}'}{k'} \times \mathbf{E}'.$$

To relate the frequency and wavenumber of the wave in \mathcal{R}' to their counterparts in \mathcal{R} , we can perform a direct transformation of the expression of the wave in \mathcal{R} . First we relate the component of \mathbf{E}'_0 to \mathbf{E}_0 and \mathbf{B}_0 components using field transformation laws. Once this step is done, we have

$$\mathbf{E}'(\mathbf{r}, t) = \mathbf{E}'_0 \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)].$$

Then we apply the Lorentz transformation $(\mathbf{r}, i\omega t) \rightarrow (\mathbf{r}', i\omega' t')$ to eliminate the variables \mathbf{r} and t in favor of \mathbf{r}' and t' . This yields

$$\mathbf{E}'(\mathbf{r}, t) = \mathbf{E}'_0 \exp[i(\mathbf{k} \cdot \{\mathbf{r}'_\perp + \gamma(\mathbf{r}'_\parallel + \beta t')\} - \omega\gamma(t' + \beta \cdot \mathbf{r}'_\parallel/c))]$$

Finally, we compare the two expressions for the wave in \mathcal{R}' . We find

$$\mathbf{k}'_\parallel = \gamma(\mathbf{k}_\parallel - \beta\omega/c), \quad \mathbf{k}'_\perp = \mathbf{k}_\perp, \quad \omega' = \gamma(\omega - \mathbf{v} \cdot \mathbf{k}_\parallel)$$

The wavevector and frequency is thus transformed by a Lorentz transformation law, allowing us to define the four-vector

$$\vec{k} = (\mathbf{k}, i\omega/c)$$

From the general property that the scalar product between two four-vectors is an invariant, we find that the phase of a plane wave

$$\phi = \vec{k} \cdot \vec{r} = \mathbf{k} \cdot \mathbf{r} - \omega t$$

is an invariant: $\phi = \phi' = \vec{k}' \cdot \vec{r}' = \mathbf{k}' \cdot \mathbf{r}' - \omega' t'$. From the property that the norm of a four-vector is an invariant, we find

$$\vec{k} \cdot \vec{k} = \mathbf{k}^2 - \frac{\omega^2}{c^2} = 0$$

that is, the dispersion relation of a plane wave in vacuum, $\omega = kc$, takes the same form in all Galilean frames.

14.7 Tensors

In order to show that Maxwell's equations take the same form in every Galilean frame, one should use the transformation laws established in previous sections for the partial derivatives, the sources (charge density and current), the fields and the potentials. This is far from obvious even if a cumbersome step by step calculation will lead to the conclusion. There is a more elegant way to demonstrate this covariance, which takes advantage of a writing of Maxwell's equation using four-tensors making their covariance fully explicit. The objective of the present section is to present Maxwell's equations in this manifestly covariant form. This requires the introduction of a four-tensors, which are multidimensional arrays obeying a transformation rule, under a Lorentz transformation of space-time coordinates.

14.7.1 Cartesian tensors

In the Euclidian space, Cartesian tensors are multidimensional arrays defined by their behavior under orthogonal transformations.

We start with a short reminder on orthogonal transformations: Consider two sets of orthogonal Cartesian unit vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and $\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3$. Since each set constitutes a complete basis of the three-dimensional space, we have

$$\mathbf{u}'_i = A_{ij} \mathbf{u}_j,$$

and the matrix A is an orthogonal matrix, that is, its transpose is equal to its inverse. The matrix A thus describes an orthogonal transformation of coordinates: it corresponds to a rotation around an axis, a reflection (mirror symmetry $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \rightarrow (-\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$) or a combination of these options. For instance an inversion $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \rightarrow (-\mathbf{u}_1, -\mathbf{u}_2, -\mathbf{u}_3)$ is a combination of reflections.

A Cartesian tensor of rank 0 is a scalar $s(\mathbf{r})$ that is invariant under a rotation of coordinates:

$$s'(\mathbf{r}') = s(\mathbf{r})$$

A Cartesian tensor of rank 1 is a vector \mathbf{V} whose three components V_i , $i = 1, 2, 3$, transform under rotations like the three components of \mathbf{r} :

$$V'_i(\mathbf{r}') = A_{ij} V_j(\mathbf{r}).$$

A Cartesian tensor of rank 2 is a nine component quantity T (a 3x3 matrix) whose components T_{km} transform under rotations by the rule

$$T'_{ij}(\mathbf{r}') = A_{ij} A_{jm} T_{km}(\mathbf{r}).$$

An example that we have already encountered in the course is the electric permittivity in the case of an anisotropic dielectric medium.

14.7.2 Lorentz tensors

Lorentz tensors are the counterparts of Cartesian tensors in Minkowsky's space: they are multidimensional arrays obeying a transformation rule under a Lorentz transformation of coordinates between two Galilean frames. This section is devoted to a short presentation of this transformation rule.

A Lorentz tensor of rank 0 is simply a scalar quantity. For instance the velocity of light, which is invariant to a change of Galilean reference frame: $c' = c$. The scalar product of two four-vectors is another example of Lorentz tensor of rank zero.

A Lorentz tensor of rank 1 is a vector whose components are transformed according to the Lorentz transformation

$$a'_\mu = L_{\mu\nu} a_\nu.$$

In other words, it is a four-vector. We have already seen several examples of four-vectors, for instance, the four-current \vec{j} or the four-potential \vec{A} .

A Lorentz tensor of rank 2 is a 4x4 matrix $T_{\mu\nu}$ whose 16 components are transformed according to the transformation rule

$$T'_{\mu\nu} = L_{\mu\alpha} L_{\nu\beta} T_{\alpha\beta}. \quad (14.8)$$

14.8 Covariant formulation of Maxwell's equations

The field-field Maxwell equations, i.e., Thomson-Maxwell's and Faraday-Maxwell's equations allowed us to define the electric and magnetic fields from the scalar and vector potentials

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}.$$

Using Cartesian axes (x, y, z) and the time-coordinate ict to build four-vectors with subscripts $1 \equiv x$, $2 \equiv y$, $3 \equiv z$, and $4 \equiv ict$, we can rewrite these field definitions as

$$\begin{cases} B_x = \partial_y A_z - \partial_z A_y = \partial_2 A_3 - \partial_3 A_2, \\ B_y = \partial_z A_x - \partial_x A_z = \partial_3 A_1 - \partial_1 A_3, \\ B_z = \partial_x A_y - \partial_y A_x = \partial_1 A_2 - \partial_2 A_1, \end{cases}$$

for the magnetic field and

$$\begin{cases} iE_x/c = \partial_{(ict)} A_x - \partial_x(iV/c) = \partial_4 A_1 - \partial_1 A_4, \\ iE_y/c = \partial_{(ict)} A_y - \partial_y(iV/c) = \partial_4 A_2 - \partial_2 A_4, \\ iE_z/c = \partial_{(ict)} A_z - \partial_z(iV/c) = \partial_4 A_3 - \partial_3 A_4, \end{cases}$$

for the electric field, where A_ν denote the four components of the four-potential

$$(A_1, A_2, A_3, A_4) = (A_x, A_y, A_z, iV/c).$$

We see that we obtain the six components of a generalized curl (in the 4d Minkowsky space), which defines the second-rank electromagnetic field-strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (14.9)$$

In terms of the original field components, the electromagnetic field strength tensor $F_{\mu\nu}$ takes the form of the 4x4 matrix

$$[\mathbf{F}] = \begin{bmatrix} 0 & B_z & -B_y & -iE_x/c \\ -B_z & 0 & B_x & -iE_y/c \\ B_y & -B_x & 0 & -iE_z/c \\ iE_x/c & iE_y/c & iE_z/c & 0 \end{bmatrix}$$

As a generalized curl, $[\mathbf{F}]$ is an antisymmetric second-rank Lorentz tensor. This means that $[\mathbf{F}]$ has only six independent components, as explicitly shown in its matrix expression. In addition, $[\mathbf{F}]$ follows the field transformation laws (14.8). This can be easily shown by expressing the components of $[F]$ as functions of the four-potential components using the definition (14.9):

$$L_{\mu\alpha} L_{\nu\beta} F_{\alpha\beta} = L_{\mu\alpha} L_{\nu\beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha).$$

Since the components of the Lorentz matrix L do not depend on coordinates, they commute with the partial derivatives with respect to space-time coordinates:

$$L_{\mu\alpha} L_{\nu\beta} F_{\alpha\beta} = L_{\mu\alpha} \partial_\alpha L_{\nu\beta} A_\beta - L_{\mu\alpha} L_{\nu\beta} \partial_\beta A_\alpha$$

For the first term on the right-hand side, we recognize the transformation rules for the four-gradient and for the four-potential:

$$L_{\mu\alpha} \partial_\alpha L_{\nu\beta} A_\beta = \partial'_\mu A'_\nu.$$

The second term on the right-hand side, we first substitute the central term $L_{\nu\beta} \partial_\beta$ using the transformation rule for the four-gradient, then we commute the resulting components of the four gradient in \mathcal{R}' with the components of the remaining Lorentz matrix, and we finally transform the four-vector, which writes

$$L_{\mu\alpha} L_{\nu\beta} \partial_\beta A_\alpha = L_{\mu\alpha} \partial'_\nu A_\alpha = \partial'_\nu L_{\mu\alpha} A_\alpha = \partial'_\nu A'_\mu.$$

In result,

$$L_{\mu\alpha} L_{\nu\beta} F_{\alpha\beta} = \partial'_\mu A'_\nu - \partial'_\nu A'_\mu = F'_{\mu\nu}.$$

The field strength tensor follows the transformation rule (14.9). It is thus a Lorentz invariant.

14.8.1 Inhomogeneous Maxwell's equations

It is straightforward to write the two inhomogeneous Maxwell equations in terms of the electromagnetic field strength tensor, using the four-current ($\mathbf{j}, ic\rho$) as a source term:

$$\partial_\nu F_{\mu\nu} = \mu_0 j_\mu. \quad (14.10)$$

On the left-hand side, the repeated ν subscript indicates a sum over the second index of the field tensor $F_{\mu\nu}$, which amounts to taking the divergence of a second-rank tensor. The result is a first-rank tensor, that is, a four-vector.

To see how Maxwell's equations are encoded in Eq. (14.10), it is instructive to expand if for the different values of μ . We retrieve Gauss-Maxwell's equation from the $\mu = 4$ component

$$\partial_\nu F_{4\nu} = \mu_0 j_4 \quad \text{and} \quad j_4 = ic\rho.$$

Writing it out, we find

$$\frac{\partial}{\partial x} \left(\frac{iE_x}{c} \right) + \frac{\partial}{\partial y} \left(\frac{iE_y}{c} \right) + \frac{\partial}{\partial z} \left(\frac{iE_z}{c} \right) + 0 = i\mu_0 c\rho,$$

which, after replacing μ_0 by $1/\epsilon_0 c^2$ is identified to Gauss-Maxwell's equation

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}.$$

The three other components for $\mu = 1, 2, 3$ are the x -, y -, and z -components of Ampère-Maxwell's equation. For instance, $\mu = 1$ yields

$$\partial_\nu F_{1\nu} = \mu_0 j_x,$$

that we expand as

$$0 + \frac{\partial}{\partial y} B_z - \frac{\partial}{\partial z} B_y + \frac{\partial}{\partial ct} \left(\frac{-iE_x}{c} \right) = \mu_0 j_x$$

and we recognize the x -component of Ampère-Maxwell's equation

$$[\nabla \times \mathbf{B}]_x - \frac{1}{c^2} \frac{\partial E_x}{\partial t} = \mu_0 j_x$$

Similarly, $\mu = 2, 3$ would give the y - and z -components of Ampère-Maxwell's equation.

Equation (14.10) is therefore a compact writing of the field-source Maxwell equations (four scalar equations). It has a much more important advantage as it shows that the field-source Maxwell equations are manifestly covariant. Indeed the left-hand side is a four-vector obtained by the contraction of the four-gradient with the field strength tensor, which are both covariant under a Lorentz transformation. The right-hand side is the four current which is covariant. Hence, Eq. (14.10) is manifestly covariant under a Lorentz transformation.

14.8.2 Field-field Maxwell's equations

The field-field Maxwell equations can also be written in terms of the field tensor $F_{\mu\nu}$. This is visible from the third-rank tensor

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0, \tag{14.11}$$

where we see that there is no repeated subscript but a cyclic permutation of the indices from one term to the next. Since the field strength tensor is antisymmetric, we have $F_{\mu\mu} = 0$ and $F_{\lambda\mu} = -F_{\mu\lambda}$, so the left-hand side of (14.11) is zero if any two indices are equal.

Now for $\lambda = 1$, $\mu = 2$, $\nu = 3$, Equation (14.11) yields

$$0 = \partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12},$$

that is, Thomson-Maxwell's equation

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = \nabla \cdot \mathbf{B} = 0.$$

The three components of Faraday-Maxwell's equation are obtained by the three remaining triads of indices. For instance $\lambda = 4$, $\mu = 1$, $\nu = 2$ yields the z -component of Faraday-Maxwell's equation:

$$\begin{aligned} 0 &= \partial_4 F_{12} + \partial_1 F_{24} + \partial_2 F_{41} \\ &= \frac{\partial B_z}{\partial c t} + \frac{\partial}{\partial x} \left(\frac{-iE_y}{c} \right) + \frac{\partial}{\partial y} \left(\frac{iE_x}{c} \right) \\ &= -i \left[\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} \right]_z \end{aligned}$$

The other choices of indices would lead to the x - and y -components of Faraday-Maxwell's equation.

14.8.3 Dual tensor

There is an alternative way to write the field-field Maxwell equations. The field strength tensor is actually not the only way to build a second-rank tensor from which Maxwell's equations can be retrieved. The field transformation laws are indeed invariant to the transformation $\mathbf{B} \rightarrow -\mathbf{E}/c$ and $\mathbf{E}/c \rightarrow \mathbf{B}$. This is called the duality transformation. Applying the duality transformation to the field strength tensor provides another second-rank tensor $G_{\mu\nu}$, which reads in matrix form

$$[\mathbf{G}] = \begin{bmatrix} 0 & -E_z/c & E_y/c & -iB_x \\ E_z/c & 0 & -E_x/c & -iB_y \\ -E_y/c & E_x/c & 0 & -iB_z \\ iB_x & iB_y & iB_z & 0 \end{bmatrix}$$

Field-field Maxwell's equations

$$\begin{cases} \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \end{cases}$$

can then be written using the dual field-strength tensor $G_{\mu\nu}$ as

$$\partial_\nu G_{\mu\nu} = 0,$$

which is a manifestly covariant equation under a Lorentz transformation, and so are the field-field Maxwell equations.

14.9 Conservation laws

To complete this chapter on covariant electrodynamics, note that conservation laws can be obtained from the symmetric, second rank, electromagnetic stress-energy tensor:

$$\Theta_{\mu\sigma} = \frac{1}{\mu_0} \left[F_{\mu\nu} F_{\sigma\nu} - \frac{1}{4} \delta_{\mu\sigma} F_{\alpha\beta} F_{\alpha\beta} \right] = \Theta_{\sigma\mu},$$

where $\delta_{\mu\sigma}$ denotes the Kronecker delta symbol. In order to relate the electromagnetic stress-energy tensor to physical quantities previously introduced in conservation laws, it is useful to give the expression for its 16 components in matrix form:

$$[\Theta] = \begin{bmatrix} -T_{xx} & -T_{xy} & -T_{xz} & i\Pi_x/c \\ -T_{yx} & -T_{yy} & -T_{yz} & i\Pi_y/c \\ -T_{zx} & -T_{zy} & -T_{zz} & i\Pi_z/c \\ i\Pi_x/c & i\Pi_y/c & i\Pi_z/c & -u_{\text{em}} \end{bmatrix},$$

where the space-space components are the negative of the components of the Maxwell stress tensor

$$T_{ij} = \epsilon_0 \left[E_i E_j + c^2 B_i B_j - \frac{1}{2} \delta_{ij} (E^2 + c^2 B^2) \right]$$

and the remaining components involve the spatial components of the Poynting vector:

$$\boldsymbol{\Pi} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}$$

and the electromagnetic energy density

$$u_{\text{em}} = \frac{\epsilon_0}{2} (\mathbf{E}^2 + c^2 \mathbf{B}^2).$$

A manifestly covariant conservation equation under Lorentz transformation writes

$$\partial_\sigma \Theta_{\sigma\mu} = -j_\nu F_{\mu\nu} = -f_\mu , \quad (14.12)$$

where the source term f_μ denotes the components of a force density four-vector defined as

$$f_\mu = (\rho \mathbf{E} + \mathbf{j} \times \mathbf{B}, i \mathbf{j} \cdot \mathbf{E}/c).$$

A careful expansion of Eq. (14.12) shows that its time component ($\mu = 4$) is just the expression for the Poynting theorem

$$-\mathbf{j} \cdot \mathbf{E} = \boldsymbol{\nabla} \cdot \boldsymbol{\Pi} + \frac{\partial u_{\text{em}}}{\partial t}.$$

The space components of Eq. (14.12) also correspond to the three components of the momentum conservation equation, which reads

$$\frac{1}{c^2} \frac{\partial \boldsymbol{\Pi}}{\partial t} = \boldsymbol{\nabla} \cdot [T] - \mathbf{f}$$

14.10 Summary

- The Lorentz transformation transforms space-time coordinates from a Galilean reference frame \mathcal{R} to another \mathcal{R}' , moving at velocity \mathbf{v} with respect to \mathcal{R} :

$$r'_\mu = L_{\mu\nu} r_\nu, \quad L = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix}$$

with

$$\beta = \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}.$$

- Four-vectors transform according to the same transformation rule as space-time coordinates under a Lorentz transformation. For instance, the four-current $\vec{j} = (\mathbf{j}, ic\rho)$, and the four-potential $\vec{A} = (\mathbf{A}, iV/c)$ are four-vectors.

- Field strength tensor: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and its dual $G_{\mu\nu}$ are expressed in matrix form as functions of the field components:

$$[\mathbf{F}] = \begin{bmatrix} 0 & B_z & -B_y & -i\frac{E_x}{c} \\ -B_z & 0 & B_x & -i\frac{E_y}{c} \\ B_y & -B_x & 0 & -i\frac{E_z}{c} \\ i\frac{E_x}{c} & i\frac{E_y}{c} & i\frac{E_z}{c} & 0 \end{bmatrix}, \quad [\mathbf{G}] = \begin{bmatrix} 0 & -\frac{E_z}{c} & \frac{E_y}{c} & -iB_x \\ \frac{E_z}{c} & 0 & -\frac{E_x}{c} & -iB_y \\ -\frac{E_y}{c} & \frac{E_x}{c} & 0 & -iB_z \\ iB_x & iB_y & iB_z & 0 \end{bmatrix}$$

- The covariant formulation of Maxwell's equations read

$$\begin{aligned} \partial_\nu F_{\mu\nu} &= \mu_0 j_\mu, \\ \partial_\nu G_{\mu\nu} &= 0, \end{aligned}$$

and means that Maxwell's equations take the same form in all Galilean reference frames.

Appendix A

Mathematical toolbox for electrodynamics

Introduction

This chapter presents mathematical theorems and relations that are used throughout the course. The goal of this chapter is to serve as reminders and as a place to find the relevant formulas. Proofs are not given in the mathematical sense but some details are provided for relations of vector calculus.

A.1 Orthogonal coordinates

- **CARTESIAN COORDINATES**

The coordinates are x, y, z . The Cartesian basis is $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$.
A line element reads

$$d\mathbf{r} = dx\mathbf{u}_x + dy\mathbf{u}_y + dz\mathbf{u}_z.$$

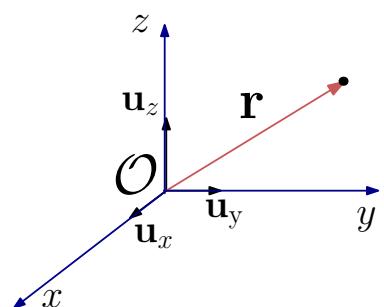
A surface element parallel to the xy -plane reads

$$dS = dx dy,$$

and permutations can be used for surfaces that are parallel to the xz - or yz -plane.

The volume element reads

$$d^3r = dx dy dz.$$



• CYLINDRICAL COORDINATES

The coordinates are r, θ, z . The local basis is $\mathbf{u}_r, \mathbf{u}_\theta, \mathbf{u}_z$, with $\mathbf{u}_\theta = \mathbf{u}_z \times \mathbf{u}_r$. A line element reads

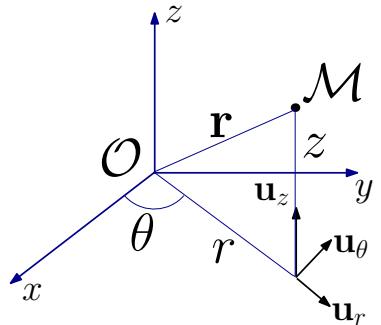
$$d\mathbf{r} = dr\mathbf{u}_r + rd\theta\mathbf{u}_\theta + dz\mathbf{u}_z.$$

A surface element writes as a product of the relevant components of the line element. For instance,

$$dS = r dr d\theta \quad \text{for a surface orthogonal to the } z\text{-axis}$$

The volume element reads

$$d^3r = r dr d\theta dz$$



• SPHERICAL COORDINATES

The coordinates are r, θ, ϕ . The local basis is $\mathbf{u}_r, \mathbf{u}_\theta, \mathbf{u}_\phi$, with $\mathbf{u}_\phi = \mathbf{u}_r \times \mathbf{u}_\theta$ and $\mathbf{u}_\theta = \mathbf{u}_\phi \times \mathbf{u}_r$. A line element reads

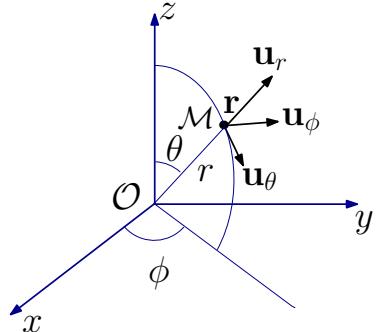
$$d\mathbf{r} = dr\mathbf{u}_r + rd\theta\mathbf{u}_\theta + r \sin \theta \mathbf{u}_\phi.$$

A surface element writes as a product of two components of the line element,

$$dS = r^2 \sin \theta d\theta d\phi$$

for a spherical surface element spanned by the θ and ϕ coordinates.
The volume element reads

$$d^3r = r^2 \sin \theta dr d\theta d\phi$$



A.2 Vectorial operators

Definition A.1: Gradient operator

Let $V(\mathbf{r}, t)$ be a scalar field measured at time t and point $\mathbf{r} \equiv x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z$, where (x, y, z) denote the Cartesian coordinates of the position vector \mathbf{r} .

The gradient operator acts on a scalar field $V(\mathbf{r}, t)$ and results in a vector with the Cartesian representation

$$\nabla V = \mathbf{u}_x \frac{\partial V}{\partial x} + \mathbf{u}_y \frac{\partial V}{\partial y} + \mathbf{u}_z \frac{\partial V}{\partial z}. \quad (\text{gradient of } V)$$

Exercise A.1 - Gradient of $1/r$

Calculate the gradient of the scalar field $V(\mathbf{r}) = \frac{1}{r}$

Solution

Writing $r = \sqrt{x^2 + y^2 + z^2}$, we can use chain's rule to differentiate

$$\frac{\partial r^{-1}}{\partial x} = -\frac{1}{2} \frac{2x}{(x^2 + y^2 + z^2)^{3/2}} = -\frac{x}{r^3}.$$

Similarly, we have $\partial V/\partial y = -y/r^3$ and $\partial V/\partial z = -z/r^3$, thus,

$$\nabla V = -\frac{x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z}{(x^2 + y^2 + z^2)^{3/2}} = -\frac{\mathbf{r}}{r^3}.$$

Definition A.2: The divergence operator

Let $\mathbf{A}(\mathbf{r}, t)$ be a vector field measured at time t and point \mathbf{r} , with the Cartesian components

$$\mathbf{A}(\mathbf{r}, t) = A_x(\mathbf{r}, t)\mathbf{u}_x + A_y(\mathbf{r}, t)\mathbf{u}_y + A_z(\mathbf{r}, t)\mathbf{u}_z$$

The divergence operator acts on a vector and results in a scalar field with the Cartesian representation

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \quad (\text{divergence of } \mathbf{A})$$

Exercise A.2 - Gradient of \mathbf{r}

Calculate the divergence of the vector field \mathbf{r}

Solution

Applying the definition in Cartesian coordinates, we find

$$\nabla \cdot \mathbf{r} = \underbrace{\frac{\partial x}{\partial x}}_1 + \underbrace{\frac{\partial y}{\partial y}}_1 + \underbrace{\frac{\partial z}{\partial z}}_1 = 3.$$

Definition A.3: The curl operator

The curl operator acts on a vector and results in a vector field with the Cartesian representation

$$\nabla \times \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \mathbf{u}_x + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \mathbf{u}_y + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \mathbf{u}_z. \quad (\text{curl of } \mathbf{A})$$

In Cartesian coordinates, applying the curl operator amounts to calculating a cross product using column vectors

$$\nabla \times \mathbf{A} = \begin{vmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{vmatrix} \times \begin{vmatrix} A_x \\ A_y \\ A_z \end{vmatrix} = \begin{vmatrix} \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \\ \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \\ \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \end{vmatrix}$$

Definition A.4: The Laplacian operator

The Laplacian of a scalar field $V(\mathbf{r}, t)$ yields a scalar field defined as the divergence of the gradient of V . Its Cartesian representation reads

$$\Delta \mathbf{A} = \nabla \cdot \nabla V = \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}. \quad (\text{Laplacian of } V)$$

The Laplacian of a vector field $\mathbf{A}(\mathbf{r}, t)$ is a vector field with the Cartesian representation

$$\Delta \mathbf{A} = \nabla^2 \mathbf{A} = \mathbf{u}_x \nabla^2 A_x + \mathbf{u}_y \nabla^2 A_y + \mathbf{u}_z \nabla^2 A_z. \quad (\text{Laplacian of } \mathbf{A})$$

Definition A.5: The \mathbf{A} dot nabla operator

or any vector field \mathbf{A} , the \mathbf{A} dot nabla operator is defined as the dot product of vector \mathbf{A} with the gradient operator. In Cartesian representation, this reads

$$(\mathbf{A} \cdot \nabla) = A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y} + A_z \frac{\partial}{\partial z}.$$

The order matters! Do not switch \mathbf{A} and ∇ in the dot product. The \mathbf{A} dot nabla operator can be applied to scalar- as well as vector-fields. In the latter case, it applies component-wise.

Exercise A.3 - $(\mathbf{A} \cdot \nabla)$ of the position vector \mathbf{r}

Calculate $(\mathbf{A} \cdot \nabla)\mathbf{r}$, where $\mathbf{A}(\mathbf{r}, t)$ denotes any vector field and \mathbf{r} denotes the position vector.

Solution

The components of the position vector are x , y and z . For the first component,

$$(\mathbf{A} \cdot \nabla)x = A_x \frac{\partial}{\partial x} x = A_x.$$

For the second and third components, we find similarly $(\mathbf{A} \cdot \nabla)y = A_y$ and $(\mathbf{A} \cdot \nabla)z = A_z$. Finally,

$$(\mathbf{A} \cdot \nabla)\mathbf{r} = \mathbf{A}$$

A.2.1 Variation of a field

The increment dV of a scalar field $V(\mathbf{r}, t)$ along a line element

$$d\mathbf{r} = dx \mathbf{u}_x + dy \mathbf{u}_y + dz \mathbf{u}_z,$$

during a time dt is given by

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz + \frac{\partial V}{\partial t}dt$$

which rewrites as

$$dV = d\mathbf{r} \cdot (\nabla V) + \frac{\partial V}{\partial t}dt = (d\mathbf{r} \cdot \nabla)V + \frac{\partial V}{\partial t}dt$$

Applying the latter result component-wise to a vector field $\mathbf{A}(\mathbf{r}, t)$, we find

$$d\mathbf{A} = (d\mathbf{r} \cdot \nabla)\mathbf{A} + \frac{\partial \mathbf{A}}{\partial t}dt$$

The first term on the right-hand-side stands for the nonuniform character of \mathbf{A} while the second term stands for its non-stationary character.

A.2.2 Vectorial operators in cylindrical coordinates

Let $V(\mathbf{r}, t) \equiv V(r, \theta, z, t)$ be a scalar field, and

$$\mathbf{A}(\mathbf{r}, t) = A_r(r, \theta, z, t)\mathbf{u}_r + A_\theta(r, \theta, z, t)\mathbf{u}_\theta + A_z(r, \theta, z, t)\mathbf{u}_z,$$

a vector field, expressed in cylindrical coordinates.

- GRADIENT

$$\nabla V = \frac{\partial V}{\partial r}\mathbf{u}_r + \frac{1}{r} \frac{\partial V}{\partial \theta}\mathbf{u}_\theta + \frac{\partial V}{\partial z}\mathbf{u}_z$$

- DIVERGENCE

$$\nabla \cdot \mathbf{A} = \frac{1}{r} \frac{\partial(rA_r)}{\partial r} + \frac{1}{r} \frac{\partial A_\theta}{\partial \theta} + \frac{\partial A_z}{\partial z}$$

- CURL

$$\nabla \times \mathbf{A} = \left[\frac{1}{r} \frac{\partial A_z}{\partial \theta} - \frac{\partial A_\theta}{\partial z} \right] \mathbf{u}_r + \left[\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right] \mathbf{u}_\theta + \frac{1}{r} \left[\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right] \mathbf{u}_z.$$

- LAPLACIAN

$$\nabla^2 V = \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial V}{\partial r} \right] + \frac{1}{r^2} \frac{\partial^2 V}{\partial \theta^2} + \frac{\partial^2 V}{\partial z^2}.$$

A.2.3 Vectorial operators in spherical coordinates

Let $V(\mathbf{r}, t) \equiv V(r, \theta, \phi, t)$ be a scalar field, and

$$\mathbf{A}(\mathbf{r}, t) = A_r(r, \theta, \phi, t)\mathbf{u}_r + A_\theta(r, \theta, \phi, t)\mathbf{u}_\theta + A_\phi(r, \theta, \phi, t)\mathbf{u}_\phi,$$

a vector field, expressed in spherical coordinates.

- GRADIENT

$$\nabla V = \frac{\partial V}{\partial r}\mathbf{u}_r + \frac{1}{r} \frac{\partial V}{\partial \theta}\mathbf{u}_\theta + \frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi}\mathbf{u}_\phi$$

- DIVERGENCE

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta A_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} \mathbf{u}_z.$$

- CURL

$$\nabla \times \mathbf{A} = \frac{1}{r \sin \theta} \left[\frac{\partial(\sin \theta A_\phi)}{\partial \theta} - \frac{\partial A_\theta}{\partial \phi} \right] \mathbf{u}_r + \frac{1}{r} \left[\frac{1}{\sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{\partial(r A_\theta)}{\partial r} \right] \mathbf{u}_\theta + \frac{1}{r} \left[\frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right] \mathbf{u}_\phi.$$

- LAPLACIAN

$$\nabla^2 V = \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial V}{\partial r} \right] + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left[\sin \theta \frac{\partial V}{\partial \theta} \right] + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 V}{\partial \phi^2}.$$

A.2.4 Vector calculus identities

In this section, $V(\mathbf{r}, t)$ and $U(\mathbf{r}, t)$ are scalar fields, $\mathbf{A}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ are vector fields.

1. The divergence of the curl of a vector is equal to zero

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

2. The curl of the gradient of a scalar is equal to zero

$$\nabla \times (\nabla V) = \mathbf{0}.$$

3. A useful identity for the physics of electromagnetic waves:

$$\nabla^2 \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}).$$

The next three identities look similar to the derivative of a product.

4 Gradient of a product

$$\nabla(UV) = U \nabla V + V \nabla U.$$

1. Curl of a product of scalar by vector field

$$\nabla \times (U\mathbf{A}) = U \nabla \times \mathbf{A} + \nabla U \times \mathbf{A}.$$

2. Divergence of a product of scalar by vector field

$$\nabla \cdot (U\mathbf{A}) = U \nabla \cdot \mathbf{A} + \nabla U \cdot \mathbf{A}.$$

The next 4 identities look *different* from the derivative of a product!

7 Gradient of a dot product

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A} \times (\nabla \times \mathbf{B}) + (\mathbf{B} \cdot \nabla)\mathbf{A} + \mathbf{B} \times (\nabla \times \mathbf{A})$$

This identity can be used to calculate the electric field for a dipole of (uniform) dipole moment \mathbf{p} : $\nabla(\mathbf{p} \cdot \mathbf{r}) = (\mathbf{p} \cdot \nabla)\mathbf{r} = \mathbf{p}$, $(\nabla \times \mathbf{r} = 0)$.

1. Curl of a cross product

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\nabla \cdot \mathbf{B})\mathbf{A} + (\mathbf{B} \cdot \nabla)\mathbf{A} - (\nabla \cdot \mathbf{A})\mathbf{B} - (\mathbf{A} \cdot \nabla)\mathbf{B}$$

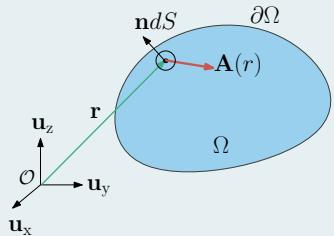
This identity can be used to calculate the magnetic field for a dipole of (uniform) dipole moment \mathbf{m} : $\nabla \times \left(\mathbf{m} \times \frac{\mathbf{r}}{r^3} \right) = -(\mathbf{m} \cdot \nabla) \frac{\mathbf{r}}{r^3} = -\nabla \left(\frac{\mathbf{m} \cdot \mathbf{r}}{r^3} \right)$

2. Divergence of a cross product

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = -\mathbf{A} \cdot \nabla \times \mathbf{B} + \mathbf{B} \cdot \nabla \times \mathbf{A}$$

A.2.5 Ostrogradski's theorem

Definition A.6: Ostrogradski's theorem



If $\mathbf{A}(\mathbf{r}, t)$ is a vector field and Ω is a volume bounded by the closed surface $\partial\Omega$, the flux of \mathbf{A} through $\partial\Omega$ is equal to the volume integral of the divergence of \mathbf{A} .

$$\oint_{\partial\Omega} \mathbf{A}(\mathbf{r}, t) \cdot d\mathbf{S}(\mathbf{r}) = \iiint_{\Omega} (\nabla \cdot \mathbf{A}(\mathbf{r}, t)) d^3r$$

Exercise A.4 - Gradient formula

For a scalar field $V(\mathbf{r}, t)$ and a volume Ω bounded by the closed surface $\partial\Omega$, demonstrate the gradient formula

$$\oint_{\partial\Omega} V(\mathbf{r}, t) d\mathbf{S}(\mathbf{r}) = \iiint_{\Omega} \nabla V(\mathbf{r}, t) d^3r.$$

Hint: Apply Ostrogradski's theorem and a suitable vector calculus formula.

Solution

Let us demonstrate the gradient formula component-wise. First project the formula on a unit vector \mathbf{u} in a Cartesian basis:

$$\mathbf{u} \cdot \oint_{\partial\Omega} V(\mathbf{r}, t) d\mathbf{S}(\mathbf{r}) = \oint_{\partial\Omega} V(\mathbf{r}, t) \mathbf{u} \cdot d\mathbf{S}(\mathbf{r}).$$

Apply Ostrogradski's theorem to the vector $V\mathbf{u}$:

$$\oint_{\partial\Omega} V(\mathbf{r}, t) \mathbf{u} \cdot d\mathbf{S}(\mathbf{r}) = \iiint_{\Omega} \nabla \cdot (V\mathbf{u}) d^3r.$$

Since $\nabla \cdot (V\mathbf{u}) = \mathbf{u} \cdot \nabla V + V \nabla \cdot \mathbf{u}$ and $\nabla \cdot \mathbf{u} = 0$, we find

$$\iiint_{\Omega} \nabla \cdot (V\mathbf{u}) d^3r = \iiint_{\Omega} \mathbf{u} \cdot \nabla V d^3r = \mathbf{u} \cdot \iiint_{\Omega} \nabla V d^3r.$$

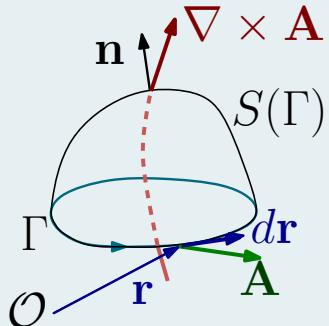
We have shown

$$\mathbf{u} \cdot \oint_{\partial\Omega} V(\mathbf{r}, t) d\mathbf{S}(\mathbf{r}) = \mathbf{u} \cdot \iiint_{\Omega} \nabla V d^3r.$$

Since this is valid for all three axes of the Cartesian basis, the gradient formula is demonstrated.

A.2.6 Stokes's theorem

Definition A.7: Stokes's theorem



Let Γ be an oriented contour and $S(\Gamma)$ be a surface bounded by Γ . The orientation of Γ implies that the surface $S(\Gamma)$ is also oriented (right-hand rule: the thumb gives the direction of the normal to the surface from the negative to the positive side when the fingers points in the same direction as for the contour orientation). The circulation of a vector field $\mathbf{A}(\mathbf{r}, t)$ around Γ is the line integral of \mathbf{A} around Γ . It is equal to the flux of $\nabla \times \mathbf{A}$ through $S(\Gamma)$.

$$\oint_{\Gamma} \mathbf{A} \cdot d\mathbf{r} = \iint_{S(\Gamma)} (\nabla \times \mathbf{A}) \cdot d\mathbf{S}$$

Exercise A.5 - Kelvin's formula

For a scalar field $V(\mathbf{r}, t)$ and an open surface $S(\Gamma)$ bounded by a closed curve Γ , demonstrate Kelvin's formula,

$$\oint_{\Gamma} V d\mathbf{r} = - \iint_{S(\Gamma)} \nabla V \times d\mathbf{S}.$$

Hint: Apply Stokes's theorem and a suitable vector calculus formula.

Solution

Let us demonstrate the formula component-wise. Taking the dot product of the left-hand-side with a unit vector \mathbf{u} of a Cartesian basis, and applying Stokes's theorem, we have

$$\mathbf{u} \cdot \oint_{\Gamma} V d\mathbf{r} = \oint_{\Gamma} V \mathbf{u} \cdot d\mathbf{r} = \iint_{S(\Gamma)} (\nabla \times V \mathbf{u}) \cdot d\mathbf{S}.$$

Since \mathbf{u} is uniform, $\nabla \times V \mathbf{u} = \nabla V \times \mathbf{u}$. Replacing in the integral and using a circular permutation in the triple product, we find

$$\iint_{S(\Gamma)} (\nabla \times V \mathbf{u}) \cdot d\mathbf{S} = \iint_{S(\Gamma)} (\nabla V \times \mathbf{u}) \cdot d\mathbf{S} = - \iint_{S(\Gamma)} (\nabla V \times d\mathbf{S}) \cdot \mathbf{u}.$$

Since this calculation is similar for all unit vectors of a Cartesian basis, Kelvin's formula is demonstrated component-wise.

Exercise A.6 - Curl formula

$$\iint_{\partial\Omega} \mathbf{A} \times d\mathbf{S} = - \iiint_{\Omega} (\nabla \times \mathbf{A}) d^3 r$$

Solution

Again, this identity is shown component-wise. Take first the dot product with a unit vector \mathbf{u} , which be pushed under the integral and then apply a circular permutation.

$$\mathbf{u} \cdot \oint_{\partial\Omega} \mathbf{A} \times d\mathbf{S} = \oint_{\partial\Omega} \mathbf{u} \cdot (\mathbf{A} \times d\mathbf{S}) = \oint_{\partial\Omega} (\mathbf{u} \times \mathbf{A}) \cdot d\mathbf{S}.$$

Apply Ostrogradsky's theorem:

$$\oint_{\partial\Omega} (\mathbf{u} \times \mathbf{A}) \cdot d\mathbf{S} = \iiint_{\Omega} \nabla \cdot (\mathbf{u} \times \mathbf{A}) d^3r.$$

From vector calculus identity 9, we have

$$\nabla \cdot (\mathbf{u} \times \mathbf{A}) = -\mathbf{u} \cdot \nabla \times \mathbf{A} + \mathbf{A} \times \underbrace{(\nabla \times \mathbf{u})}_{\mathbf{0}}.$$

Thus

$$\iiint_{\Omega} \nabla \cdot (\mathbf{u} \times \mathbf{A}) d^3r = \iiint_{\Omega} -\mathbf{u} \cdot (\nabla \times \mathbf{A}) d^3r = -\mathbf{u} \cdot \iiint_{\Omega} (\nabla \times \mathbf{A}) d^3r.$$

We have shown for any unit vector

$$\mathbf{u} \cdot \oint_{\partial\Omega} \mathbf{A} \times d\mathbf{S} = -\mathbf{u} \cdot \iiint_{\Omega} (\nabla \times \mathbf{A}) d^3r,$$

which completes the demonstration of the curl formula.

Appendix B

The Dirac distribution

The goal of this section is to introduce the *Dirac distribution*, that is used as a density function for discrete charge distributions

Let us first restrict ourselves to one dimension and consider the Dirac distribution δ centered at $x = 0$. Strictly speaking, δ is defined as a linear application over the space \mathcal{D} of *test functions*, that is, functions which are infinitely differentiable and are non-zero inside a bounded interval. The Dirac distribution is defined as

$$(\delta, \varphi) = \varphi(0) \in \mathbb{C} \quad \varphi \in \mathcal{D} \quad (\text{B.1})$$

and as such the δ is not a function, but a linear functional over \mathcal{D} . It does not make any sense, for example, trying to evaluate δ at a particular value of x . Often, the notation $\delta(x)$ is used to designate the Dirac distribution centered at $x = 0$, whereas $\delta(x - x_0)$ denotes the Dirac centered at $x = x_0$, which is defined as

$$(\delta_{x_0}, \varphi) = \varphi(x_0) \in \mathbb{C} \quad \varphi \in \mathcal{D}$$

B.1 Functions that tend to the Dirac distribution

Consider for example the following sequence of functions:

$$\delta_n(x) = \begin{cases} \frac{n}{2} & |x| < 1/n, \\ 0 & |x| > 1/n. \end{cases} \quad (\text{B.2})$$

For all $n \in \mathbb{N}$ the area under the curve of δ_n is equal to 1. Then, note that as n increases, the function δ_n describes a rectangle that is increasingly concentrated around the origin. This sequence tends to zero for all $x \neq 0$, that is, it converges to the null function at all points except for $x = 0$. In consequence $\int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) dx = 0$. However, for any test function $\varphi \in \mathcal{D}$, one has

$$\int_{\mathbb{R}} \varphi(x) \delta_n(x) dx = \frac{n}{2} \int_{-1/n}^{1/n} \varphi(x) dx = \varphi(0)$$

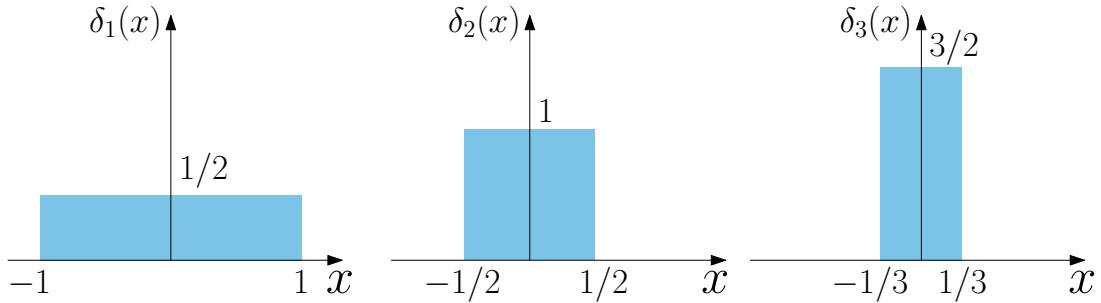


Figure B.1: First three functions given by the sequence δ_n

with $-1 \leq \theta_n \leq 1$ (theorem of intermediate values). Taking the limit when $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \varphi(x) \delta_n(x) dx = \lim_{n \rightarrow \infty} \varphi(\theta_n/n) = \varphi(0) = (\delta, \varphi)$$

The sequence of functions δ_n has therefore the desired property (B.1) of the Dirac δ as $n \rightarrow \infty$. By defining the linear functional

$$(\delta_n, \varphi) = \int_{\mathbb{R}} \delta_n(x) \varphi(x) dx,$$

we see that the sequence $(\delta_n)_{n \in \mathbb{N}}$ converges to δ in the sense of distributions:

$$\lim_{n \rightarrow \infty} (\delta_n, \varphi) = (\delta, \varphi) = \varphi(0).$$

There is in fact an infinite number of sequences that converge to δ , for example a sequence of Gaussian functions of the form:

$$h_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}$$

converges to δ as well, since $\lim_{n \rightarrow \infty} \int_{\mathbb{R}} h_n(x) \varphi(x) dx = \varphi(0)$ for every test function φ .

Note Throughout the text, we will recurrently use the following abuse of notation

$$\int_{\mathbb{R}} \delta(x) \varphi(x) dx = \varphi(0) \tag{B.3}$$

even if this integral is not really defined, since δ is not a function. Every time we write this, we can think of it as a limit of integrals of the form $\int_{\mathbb{R}} \delta(x) \varphi(x) dx \equiv \lim_{n \rightarrow \infty} \int_{\mathbb{R}} h_n(x) \varphi(x) dx = \varphi(0)$, where h_n is any sequence of functions that converges to δ in the sense stated above. For example, with the above abuse of notation, we write

$$\iiint_{\mathbb{R}^3} \delta(\mathbf{r}') \frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} d^3 r' = \frac{\mathbf{r}}{|\mathbf{r}|^3} \tag{B.4}$$

B.2 Solution of Poisson's equation in a finite volume

Let us consider a region in space $\Omega \subseteq \mathbb{R}^3$. It is possible to write the potential as an integral over Ω and over its boundary $\partial\Omega$. This makes use of Green's theorem and Green's identity that will be demonstrated below.

B.2.1 Green's theorem

Let us write a vector field \mathbf{A} under the form

$$\mathbf{A} = \phi \nabla \psi$$

where ϕ and ψ are two scalar fields. Then

$$\nabla \cdot \mathbf{A} = \nabla \cdot (\phi \nabla \psi) = \nabla \phi \cdot \nabla \psi + \phi \nabla^2 \psi$$

and using Green-Ostrogradsky's theorem, we obtain *Green's first identity*

$$\iiint_{\Omega} \underbrace{(\nabla \phi \cdot \nabla \psi + \phi \nabla^2 \psi)}_{\nabla \cdot \mathbf{A}} d^3x = \iint_{\partial\Omega} \underbrace{\phi \nabla \psi}_{\mathbf{A}} \cdot d\mathbf{S} \quad (\text{B.5})$$

Now, we can exchange the roles of ψ and ϕ to obtain an equivalent identity

$$\iiint_{\Omega} \underbrace{(\nabla \phi \cdot \nabla \psi + \psi \nabla^2 \phi)}_{\nabla \cdot \mathbf{A}} d^3x = \iint_{\partial\Omega} \psi \nabla \phi \cdot d\mathbf{S}$$

and subtracting the latter equation from Eq. (B.5), we obtain *Green's theorem*

$$\iiint_{\Omega} (\phi \nabla^2 \psi - \psi \nabla^2 \phi) d^3x = \iint_{\partial\Omega} \{\phi \nabla \psi - \psi \nabla \phi\} \cdot d\mathbf{S} \quad (\text{B.6})$$

Definition B.1: Poisson's integral equation

If V satisfies Poisson's equation in Ω , then, for $\mathbf{x} \in \Omega$:

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iiint_{\Omega} \frac{\rho(\mathbf{x}') d^3x'}{|\mathbf{x} - \mathbf{x}'|} - \frac{1}{4\pi} \iint_{\partial\Omega} \frac{d\mathbf{S}(\mathbf{x}') \cdot \mathbf{E}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} + \frac{1}{4\pi} \iint_{\partial\Omega} d\mathbf{S}(\mathbf{x}') \cdot \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} V(\mathbf{x}')$$

In consequence, the potential at any point $\mathbf{x} \in \Omega$ can be written in terms of the charge density in Ω , and of the value of the potential and the electric field at the boundary $\partial\Omega$. All the information about the exterior of Ω is therefore contained by V and \mathbf{E} in its boundary $\partial\Omega$.

Proof

We use Green's theorem B.6 with $\phi(\mathbf{x}') = V(\mathbf{x}')$ and $\psi(\mathbf{x}') = 1/|\mathbf{x} - \mathbf{x}'|$. Recalling the fundamental identity

$$\nabla'^2\psi(\mathbf{x}') = -4\pi\delta(\mathbf{x} - \mathbf{x}')$$

and that $\nabla^2V = \rho/\epsilon_0$, then

$$\begin{aligned} \iiint_{\Omega} \left\{ V(\mathbf{x}') \nabla'^2\psi(\mathbf{x}') - \psi(\mathbf{x}') \nabla'^2V(\mathbf{x}') \right\} d^3x' \\ = -4\pi \iiint_{\Omega} \left\{ V(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}') - \frac{\rho(\mathbf{x}')}{4\pi\epsilon_0 |\mathbf{x} - \mathbf{x}'|} \right\} d^3x' \\ = \oint_{\partial\Omega} \left\{ V(\mathbf{x}') \nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|} \cdot d\mathbf{S} - \frac{1}{|\mathbf{x} - \mathbf{x}'|} \nabla' V(\mathbf{x}') \cdot d\mathbf{S} \right\} \end{aligned}$$

Considering $\mathbf{x} \in \Omega$ and that $\nabla' \frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3}$, one obtains Poisson's integral equation.

Corollary

By taking $\Omega = \mathbb{R}^3$, and for an integrable charge density ($\iiint_{\mathbb{R}^3} |\rho(\mathbf{x}')| d^3x' < \infty$), we obtain the known result,

$$V(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x'$$

B.2.2 Uniqueness of the solution to the Poisson-Dirichlet problem

Poisson-Dirichlet's problem reads

$$\begin{cases} \mathbf{x} \in \Omega : \quad \nabla^2V(\mathbf{x}) = -\frac{\rho(\mathbf{x})}{\epsilon_0} \\ \mathbf{x} \in \partial\Omega : \quad V(\mathbf{x}) = V_D(\mathbf{x}) \end{cases}$$

To show that the solution is unique, let $V_1(x)$ and $V_2(x)$ be two solutions of this problem. Then

$$\psi(\mathbf{x}) = V_2(\mathbf{x}) - V_1(\mathbf{x})$$

is a solution to Laplace's equation with homogeneous Dirichlet's boundary conditions, that is

$$\begin{cases} \mathbf{x} \in \Omega : \quad \nabla^2\psi(\mathbf{x}) = 0 \\ \mathbf{x} \in \partial\Omega : \quad \psi(\mathbf{x}) = 0 \end{cases}$$

and from Green's first identity (B.5)

$$\iiint_{\Omega} \left\{ \underbrace{\nabla\psi \cdot \nabla\psi}_{|\nabla\psi|^2} + \psi \underbrace{\nabla^2\psi}_{=0} \right\} d^3x = \oint_{\partial\Omega} \psi \nabla\psi \cdot \mathbf{n} dS = 0$$

but since $|\nabla\psi| \geq 0$, we conclude that

$$\nabla\psi(\mathbf{x}) = 0 \quad \text{for } \mathbf{x} \in \Omega.$$

Then $\psi(\mathbf{x})$ is constant for $\mathbf{x} \in \Omega$, and since $\psi = 0$ on $\partial\Omega$, the value of this constant is zero. Finally

$$\psi(\mathbf{x}) = V_2(\mathbf{x}) - V_1(\mathbf{x}) = 0 \quad \text{for } \mathbf{x} \in \Omega$$

and $V_1 = V_2$ in Ω , which demonstrates that the solution of the Poisson-Dirichlet problem is unique.

B.2.3 Uniqueness of the solution to the Poisson-Neumann problem

The Poisson-Neumann problem reads

$$\begin{cases} \mathbf{x} \in \Omega : & \nabla^2 V(\mathbf{x}) = -\frac{\rho(\mathbf{x})}{\epsilon_0} \\ \mathbf{x} \in \partial\Omega : & \nabla V(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = -E_N(\mathbf{x}) \end{cases}$$

Suppose that $V_1(\mathbf{x})$ and $V_2(\mathbf{x})$ are two solutions of this problem, and define

$$\psi(\mathbf{x}) = V_2(\mathbf{x}) - V_1(\mathbf{x}).$$

It follows that $\psi(\mathbf{x})$ satisfies Laplace's equation with a homogeneous Neumann boundary condition

$$\begin{cases} \mathbf{x} \in \Omega : & \nabla^2 \psi(\mathbf{x}) = 0, \\ \mathbf{x} \in \partial\Omega : & \nabla \psi(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0. \end{cases}$$

Green's first identity (B.5) gives

$$\iiint_{\Omega} |\nabla \psi|^2 d^3x = \iint_{\partial\Omega} \psi \underbrace{\nabla \psi \cdot \mathbf{n}}_{=0} dS = 0$$

and we conclude $\nabla\psi = \mathbf{0}$, so $\psi = V_1 - V_2$ is constant in Ω . The two solutions V_1 and V_2 are equal up to an arbitrary constant which does not modify the resulting electric field. In this sense, the solution to Poisson-Neumann's problem is also unique.

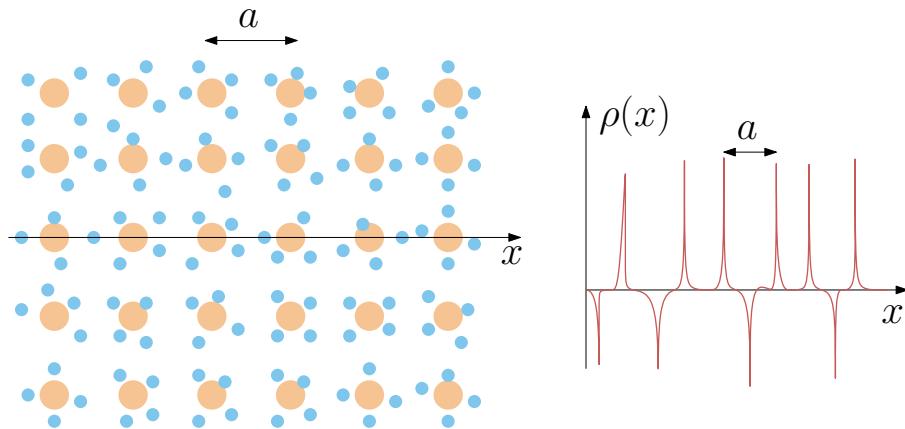
Appendix C

From discrete to continuous matter

Since any piece of matter is composed of elementary particles separated by vacuum, the charge densities and the electric field have singularities at the atomic level that can only be mathematically treated with the Dirac distribution. The case of continuous charge distributions seen in 1 consists in spatially averaging both the charge density and the resulting electric field.

C.1 Spatial averaging

Let us consider a point charge q at position \mathbf{r}_0 . Its charge density can be written in terms of a Dirac distribution $\rho_{\text{micro}} = q\delta(\mathbf{r} - \mathbf{r}_0)$. Since any elementary brick of matter (atoms and molecules) is composed of discrete elementary particles, any charge density will be in general a sum of Dirac distributions. The charge density in matter thus varies very rapidly, at the scale of $a = 0.1 \text{ nm}$.

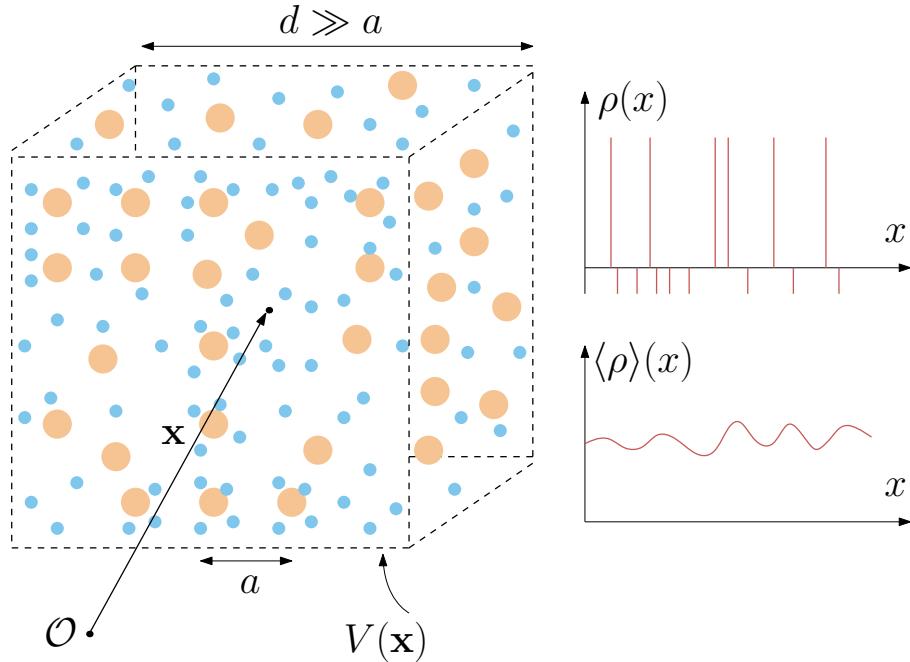


In many cases, we are not interested in these rapid variations but instead would like to work with an averaged charge density. This is the case for example when dealing with macroscopic objects, for which we only care about variations of the charge density and of the electric field over a length scale $d \gg a$. For this, we define the macroscopic charge density and electric field at every point \mathbf{r} as the average over a volume $V(\mathbf{r}) \sim d^3$ centered around \mathbf{r}

$$\langle \rho \rangle(\mathbf{r}) = \frac{1}{V} \iiint_{V(\mathbf{r})} \rho(\mathbf{r}') d^3 r'$$

$$\langle \mathbf{E} \rangle(\mathbf{r}) = \frac{1}{V} \iiint_{V(\mathbf{r})} \mathbf{E}(\mathbf{r}') d^3 r'$$

they are therefore smooth, continuous functions of \mathbf{r} provided that $d \gg a$.



This averaging may be written in a more convenient way. If we define the function φ as

$$\varphi(\mathbf{r}) = \begin{cases} \frac{1}{V} & \text{if } \mathbf{r} \in V(\mathbf{r}) \\ 0 & \text{otherwise} \end{cases}$$

then $\varphi(\mathbf{r} - \mathbf{r}')$ equals to $1/V$ in a volume V centered at \mathbf{r}' and 0 everywhere else, and we can write

$$\langle \rho \rangle(\mathbf{r}) = \iiint_{\mathbb{R}^3} \rho(\mathbf{r}') \varphi(\mathbf{r} - \mathbf{r}') d^3 r' = \rho * \varphi \quad (\text{C.1})$$

$$\langle \mathbf{E} \rangle(\mathbf{r}) = \iiint_{\mathbb{R}^3} \mathbf{E}(\mathbf{r}') \varphi(\mathbf{r} - \mathbf{r}') d^3 r' = \mathbf{E} * \varphi \quad (\text{C.2})$$

where $*$ stands for the convolution product. For example, a point charge q located at \mathbf{r}_i defines a charge distribution $\rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{r}_i)$. The associated averaged charge distribution then writes

$$\langle \rho \rangle(\mathbf{r}) = \iint_{\mathbb{R}^3} q\delta(\mathbf{r}' - \mathbf{r}_i) \varphi(\mathbf{r} - \mathbf{r}') d^3 r' = q\varphi(\mathbf{r} - \mathbf{r}_i)$$

which represents a charge q uniformly distributed over a volume V centered around \mathbf{r}_i . For a system of N charges q_i located at positions \mathbf{r}_i , with $i = 1, 2, \dots, N$, the average charge density then reads

$$\langle \rho \rangle(\mathbf{r}) = \sum_{i=1}^N q_i \varphi(\mathbf{r} - \mathbf{r}_i).$$

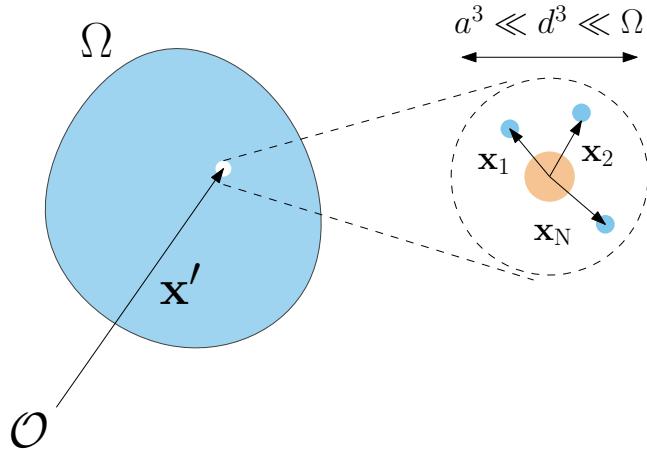
in general, the spatial average over the scale d of any quantity f is defined as

$$\langle f(\mathbf{r}) \rangle = \iiint_{\mathbb{R}^3} f(\mathbf{r}') \varphi(\mathbf{r} - \mathbf{r}') d^3 r' = f * \varphi$$

(C.3)

C.2 Continuous charge distributions

Consider now a volume Ω representing a macroscopic object, and let us consider one particular atom whose center of mass is at position \mathbf{r}' . It is composed of N charges q_i located at \mathbf{r}_i with respect to \mathbf{r}' , with $i = 1, 2, \dots, N$.



The microscopic charge density associated to this atom is $\rho_{\text{at}}(\mathbf{r}) = \sum_{i=1}^N q_i \delta(\mathbf{r} - \mathbf{r}' - \mathbf{r}_i)$. The average charge density then reads

$$\langle \rho_{\text{at}} \rangle(\mathbf{r}) = \sum_{i=1}^N q_i \varphi(\mathbf{r} - \mathbf{r}' - \mathbf{r}_i)$$

with d the characteristic extension of the function φ in each direction around $\mathbf{0}$. Now, if $d \gg a$, where $a \sim 0.1 \text{ nm}$ is the typical size of an atom, the function φ is slowly varying in the above sum, so that to the lowest order, $\varphi(\mathbf{r} - \mathbf{r}' - \mathbf{r}_i) \approx \varphi(\mathbf{r} - \mathbf{r}')$ and

$$\langle \rho_{\text{at}}(\mathbf{r}) \rangle \approx \sum_{i=1}^N q_i \varphi(\mathbf{r} - \mathbf{r}') = Q_{\text{at}} \varphi(\mathbf{r} - \mathbf{r}'),$$

where $Q_{\text{at}} = \sum_{i=1}^N q_i$ is simply the total charge contained by the atom. On the other hand, the electric field at position \mathbf{r} generated by this atom is given by

$$\mathbf{E}_{\text{at}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^N \frac{q_i(\mathbf{r} - \mathbf{r}' - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}' - \mathbf{r}_i|^3}$$

and by neglecting $|\mathbf{r}' - \mathbf{r}_i|$ with respect to $|\mathbf{r} - \mathbf{r}'|$, we obtain the electric field of a point charge Q_{at} at \mathbf{r}'

$$\mathbf{E}_{\text{at}}(\mathbf{r}) \approx \frac{Q_{\text{at}}}{4\pi\epsilon_0} \frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}$$

and the average electric field writes

$$\langle \mathbf{E}_{\text{at}}(\mathbf{r}) \rangle = \iiint_{\mathbb{R}^3} \mathbf{E}_{\text{at}}(\mathbf{u}) \varphi(\mathbf{r} - \mathbf{u}) d^3 u \approx \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{Q_{\text{at}} \varphi(\mathbf{r} - \mathbf{u})(\mathbf{u} - \mathbf{r}')}{|\mathbf{u} - \mathbf{r}'|^3} d^3 u$$

by changing variables $\mathbf{v}' = \mathbf{r}' + \mathbf{r} - \mathbf{u}$

$$\langle \mathbf{E}_{\text{at}}(\mathbf{r}) \rangle \approx \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{Q_{\text{at}} \varphi(\mathbf{v}' - \mathbf{r}')(\mathbf{r} - \mathbf{v}')}{|\mathbf{r} - \mathbf{v}'|^3} d^3 v' = \frac{1}{4\pi\epsilon_0} \iiint_{\mathbb{R}^3} \frac{\langle \rho_{\text{at}} \rangle(\mathbf{v}')(\mathbf{r} - \mathbf{v}')}{|\mathbf{r} - \mathbf{v}'|^3} d^3 v'$$

so that the averaged electric field created by the atom is given by a Coulomb integral where the charge density ρ_{at} is replaced by the averaged charge density of the atom. Finally, since the volume Ω is a collection of M atoms located at \mathbf{r}'_k with $k = 1, 2, \dots, M$, one defines the continuous charge distribution ρ at every point of Ω as

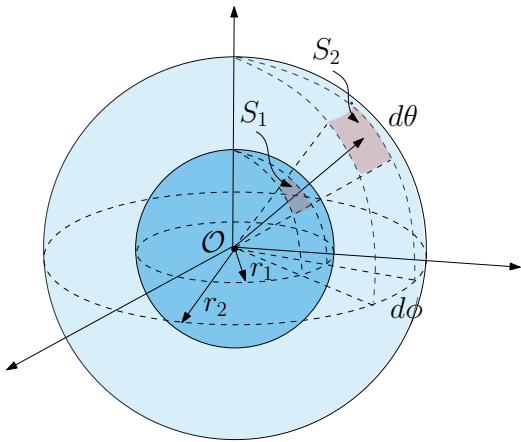
$$\rho(\mathbf{r}) = \sum_{k=1}^M Q_k \varphi(\mathbf{r} - \mathbf{r}'_k)$$

and the average electric field is obtained by calculating the Coulomb integral with the above average charge distribution.

Appendix D

The solid angle

The notion of solid angle corresponds to a two-dimensional angle representing the apparent size of a surface seen from a point \mathcal{O} . For example, consider two concentric spheres of radii r_1 and r_2 respectively, as shown in the figure below.



an arbitrary surface element $d\mathbf{S} = dS\mathbf{n}$ (not necessarily a portion of a sphere) at distance r from \mathcal{O} , the solid angle $d\Omega$ that it subtends will be the same as that of its projection on a sphere, as shown in the figure. If this sphere has a unit radius, then the solid angle will be equal to the projected surface of $d\mathbf{S}$ on the sphere, i.e. $d\Omega = (1/r^2)d\mathbf{S} \cdot \mathbf{u}_r$.

Summarizing, the solid angle Ω subtended by S at \mathcal{O} is that part of the area of a sphere of unit radius centered at \mathcal{O} that is cut off by the rays fanning out radially from \mathcal{O} to every point of S .

To calculate a solid angle:

- place the origin at the observation point: \mathcal{O}
- Use a spherical basis. Project $d\mathbf{S} = dS\mathbf{n}$ onto \mathbf{u}_r and divide by the distance squared.

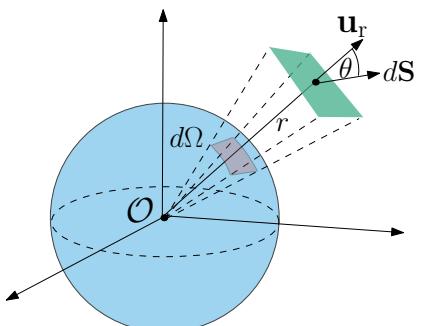
$$\Omega = \iint_S d\Omega(\mathbf{r}) = \iint_S d\mathbf{S} \cdot \frac{\mathbf{u}_r}{r^2}$$

For a surface S that encloses a volume V : $\Omega = \begin{cases} 4\pi, & \text{if } \mathcal{O} \in V \\ 0, & \text{if } \mathcal{O} \notin V \end{cases}$

The two spherical portions S_1 and S_2 , seen from \mathcal{O} , have the same apparent size. This is because both subtend the same solid angle. Mathematically, a surface element in spherical coordinates is written $dS = r^2 \sin \theta d\theta d\phi$. The elementary solid angle is defined as $d\Omega = \sin \theta d\theta d\phi$. It is easy to see that the two surfaces of the figure are described by the same intervals of θ and of ϕ , and then both subtend the same total solid angle $\Omega = \int d\Omega$. The surfaces S_1 and S_2 then read:

$$S_1 = r_1^2 \Omega, \quad S_2 = r_2^2 \Omega.$$

The solid angle Ω is measured in steradians, and can vary between 0 and 4π . For the more general case of



Particular case: S is a sphere centered at \mathcal{O}

$$d\Omega = dS \frac{\mathbf{u}_r \cdot \mathbf{n}}{r^2} = \frac{dS}{r^2} = \sin \theta d\theta d\phi$$

Solid angles are used for instance to demonstrate Gauss's law by purely geometrical arguments.

Appendix E

Legendre polynomials

The main properties of Legendre polynomials that are useful in electrostatics are listed below.

- For $l \in \mathbb{N}_0$, the Legendre polynomial P_l is a polynomial of order l , which is a solution of Legendre's differential equation

$$\frac{d}{dx} \left\{ (1 - x^2) \frac{dP_l(x)}{dx} \right\} + l(l+1)P_l(x) = 0 \quad (\text{E.1})$$

for $|x| \leq 1$. P_l can be explicitly computed by Rodrigues formula:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} [(x^2 - 1)^l] \quad (\text{E.2})$$

with $P_l(1) = 1$ and $P_l(-1) = (-1)^l$ for every $l \in \mathbb{N}_0$. The first Legendre polynomials are

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2} (3x^2 - 1) \end{aligned}$$

and satisfy the recursive formula

$$(l+1)P_{l+1}(x) = (2l+1)xP_l(x) - lP_{l-1}(x) \quad (\text{E.3})$$

- The Legendre polynomials are such that

$$P_l(-x) = (-1)^l P_l(x)$$

and so if l is even (resp. odd), P_l is an even function (resp. odd function).

- One has the identity

$$P_l(x) = \frac{1}{2l+1} \frac{d}{dx} [P_{l+1}(x) - P_{l-1}(x)]. \quad (\text{E.4})$$

4. The set of Legendre polynomials $\{P_l\}_{l \in \mathbb{N}_0}$ is orthogonal in $[-1, 1]$. This means:

$$\int_{-1}^1 P_{l'}(x) P_l(x) dx = \begin{cases} 0 & \text{if } l \neq l' \\ \frac{2}{(2l+1)} & \text{if } l = l' \end{cases} \quad (\text{E.5})$$

5. The set of Legendre polynomials $\{P_l\}_{l \in \mathbb{N}_0}$ is a complete orthogonal set in $[-1, 1]$. In consequence, if $f : [-1, 1] \mapsto \mathbb{C}$ is such that $\int_{[-1,1]} |f(x)|^2 dx < \infty$, then it admits an expansion in Legendre polynomials

$$f(x) = \sum_{l=0}^{\infty} A_l P_l(x) \quad (\text{E.6})$$

where the coefficients A_l are given by

$$A_l = \frac{2l+1}{2} \int_{-1}^1 P_l(x) f(x) dx \quad (\text{E.7})$$

6. We will admit the following equality

$$\frac{1}{\sqrt{1 - 2xr + r^2}} = \sum_{l=0}^{\infty} P_l(x) r^l \quad r < 1 \quad (\text{E.8})$$

which allows, in electrostatics, to expand $1/|\mathbf{x} - \mathbf{x}'|$ in terms of Legendre polynomials. Indeed, considering that, in spherical coordinates, $|\mathbf{x}| = r$, $|\mathbf{x}'| = r'$ and $\mathbf{x} \cdot \mathbf{x}' = rr' \cos \theta$, then:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}}.$$

Assuming, without loss of generality, that $r > r'$:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{r \sqrt{1 + (r'/r)^2 - 2(r'/r) \cos \theta}}$$

and comparing with the identity (E.8), we obtain

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \sum_{l=0}^{\infty} P_l(\cos \theta) \frac{r'^l}{r^{l+1}} \quad \text{for } r > r'. \quad (\text{E.9})$$

Appendix F

Spherical harmonics

Spherical harmonics $Y_{lm}(\theta, \phi)$, where l and m are integers and θ and ϕ denote the azimuthal and polar angles of spherical coordinates, are eigenfunctions of the eigenvalue problem defined by the partial differential equation

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} = -l(l+1)Y.$$

Spherical harmonics $Y_{lm}(\theta, \phi)$ are complex-valued. For $|m| \leq l$, the spherical harmonics are expressed in terms of the associated Legendre polynomials,

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} \frac{(l-m)!}{(l+m)!} P_l^m(\cos \theta) e^{im\phi}, \quad \text{for } m \geq 0,$$

and $Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}^*(\theta, \phi)$.

The associated Legendre polynomials are expressed in terms of Legendre's polynomials as

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x).$$

The spherical harmonics form a complete and orthonormal basis satisfying

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{lm}(\theta, \phi) Y_{l'm'}^*(\theta, \phi) = \delta_{ll'} \delta_{mm'} \\ \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \delta(\cos \theta - \cos \theta') \delta(\phi - \phi').$$

The first few spherical harmonics are, in spherical coordinates, or as functions of Cartesian coordinates $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, $z = r \cos \theta$,

$$Y_{00}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}, \\ Y_{10}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos \theta = \sqrt{\frac{3}{4\pi}} \frac{z}{r} \\ Y_{1\pm 1}(\theta, \phi) = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} = \mp \sqrt{\frac{3}{8\pi}} \frac{x \pm iy}{r}.$$

Spherical harmonics addition theorem: For $\mathbf{r} = (r, \theta, \phi)$ and $\mathbf{r}' = (r', \theta', \phi')$ referring to a coordinate system with the same origin,

$$P_l(\mathbf{u}_r \cdot \mathbf{u}_{r'}) = \frac{4\pi}{2l+1} \sum_{l=-l}^l Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi).$$

Spherical harmonics transform under space inversion as

$$Y_{lm}(-\mathbf{u}_r) = (-1)^l Y_{lm}(\mathbf{u}_r).$$