

Lectures on Quantum Mechanics

(nonlinear PDE point of view)

A.I.Komech¹

Department of Mechanics-Mathematics

Moscow State University

Moscow 119992, Russia

Abstract

We expose the Schrödinger quantum mechanics with traditional applications to Hydrogen atom: the calculation of the atom spectrum via the Schrödinger, Pauli and Dirac equations, the Heisenberg representation, the selection rules, the calculation of quantum and classical scattering of light (Thomson's cross section), photoeffect (Sommerfeld's cross section), quantum and classical scattering of electrons (Rutherford's cross section), normal and anomalous Zeemann effect (Landé's factor), polarization and dispersion (Kramers-Kronig's formula), diamagnetic susceptibility (Langevin's formula).

We give a detailed discussion of the experimental and theoretical background for the introduction of the Schrödinger, Pauli, and Dirac equations, as well as for the introduction of the Maxwell equations. We explain in detail all basic theoretical concepts: the introduction of the quantum stationary states, charge density and electric current density, quantum magnetic moment, electron spin and spin-orbital coupling in "vector model" and in the Russel-Saunders approximation, differential cross section of scattering, the Lorentz theory of polarization and magnetization, the Einstein special relativity and covariance of the Maxwell Electrodynamics.

We explain all the details of the calculations and mathematical tools: Lagrangian and Hamiltonian formalism for the systems with finite degree of freedom and for fields, Geometric Optics, the Hamilton-Jacobi equation and WKB approximation, Noether theory of invariants including the theorem on currents, four conservation laws (energy, momentum, angular momentum and charge), Lie algebra of angular momentum and spherical functions, scattering theory (limiting amplitude principle and limiting absorption principle), the Lienard-Wiechert formulas, Lorentz group and Lorentz formulas, Pauli theorem and relativistic covariance of the Dirac equation, etc.

We give a detailed overview of the conceptual development of the quantum mechanics, and expose main achievements of the "old quantum mechanics" in the form of exercises.

One of our basic aims in writing this book is an open and concrete discussion of the problem of mathematical description of the fundamental quantum phenomena: (i) Bohr's quantum transitions and (ii) de Broglie's wave-particle duality. Both phenomena cannot be described by autonomous linear dynamical equations. We treat these two problems using recent progress in the theory of global attractors of nonlinear hyperbolic PDEs. Namely, we suggest that (i) in the presence of an external confining potential, the quantum stationary states form a global attractor of the coupled Maxwell-Schrödinger or Maxwell-Dirac equations, and (ii) the wave-particle duality corresponds to the soliton-like asymptotics for the solutions of the translation-invariant coupled equations without an external potential.

We emphasize, in the whole exposition, that the coupled equations are nonlinear, and it is this nonlinearity that lies behind all traditional perturbative calculations known as the *Born approximation*. We suggest that both fundamental quantum phenomena could be described by this nonlinear coupling. The suggestion is confirmed by recent results on the global attractors and soliton asymptotics for model nonlinear hyperbolic PDEs.

¹Supported partly by the Faculty of Mathematics of the Vienna University, Max-Planck Institute for Mathematics in the Sciences (Leipzig), Project ONDES (INRIA, Rocquencourt), and Wolfgang Pauli Institute.

Contents

0	Preface	8
1	Introduction: Quantum Chronology 1859-1927	15
1.1	Missing “Matter Equation”	15
1.2	Thermodynamics, Optics and Electrodynamics	15
1.3	Atomic Physics	17
I	Lagrangian Field Theory	23
2	Euler-Lagrange Field Equations	25
2.1	Klein-Gordon and Schrödinger Equations	25
2.2	Lagrangian Density	25
2.3	Free Equations	26
2.4	The Equations with Maxwell Field	27
2.5	Action Functional	27
2.6	Hamilton Least Action Principle	28
3	Four Conservation Laws for Lagrangian Fields	29
3.1	Time Invariance: Energy Conservation	29
3.2	Translation Invariance: Momentum Conservation	31
3.3	Rotation Invariance: Angular Momentum Conservation	31
3.4	Phase Invariance: Charge Conservation	32
4	Lagrangian Theory for the Maxwell Field	34
4.1	Maxwell Equations and Potentials. Lagrangian Density	34
4.2	Lagrangian for Charged Particle in Maxwell Field	36
4.3	Hamiltonian for Charged Particle in Maxwell Field	38
II	Schrödinger Equation	39
5	Geometric Optics and Schrödinger Equation	41
5.1	Straight Line Propagation for the Free Equations	41
5.2	WKB Asymptotics for Schrödinger Equation with a Maxwell Field	44
6	Schrödinger Equation and Heisenberg Representation	47
6.1	Electrons and Cathode Rays	47
6.2	Quantum Stationary States	47
6.3	Four Conservation Laws for Schrödinger Equation	48
6.4	Quantum Observables and Heisenberg Representation	49
7	Coupling to the Maxwell Equations	52
7.1	Charge and Current Densities and Gauge Invariance	52
7.2	Electron Beams and Heisenberg’s Uncertainty Principle	53
7.3	Quantum Stationary States and Attractors	54

7.4	Charge Continuity Equations	56
7.5	Born Approximation	57
III	Application of Schrödinger Theory	59
8	Spectrum of Hydrogen Atom	61
8.1	Spherical Symmetry and Separation of Variables	61
8.2	Spherical Laplacian	63
8.3	Radial Equation	64
9	Spherical Spectral Problem	66
9.1	Hilbert-Schmidt Argument	66
9.2	Lie Algebra of Angular Momenta	66
9.3	Irreducible Representations of Commutation Relations	67
9.4	Spherical Harmonics	68
9.5	Angular Momenta in Spherical Coordinates	69
10	Atom Dipole Radiation and Selection Rules	70
10.1	Rydberg-Ritz Combination Principle	70
10.2	Selection Rules for Cylindrical Symmetry	72
10.3	Selection Rules for Orbital Momentum	74
11	Classical Scattering of Light: Thomson formula	77
11.1	Incident Plane Wave	77
11.2	Scattering Problem	77
11.3	Neglecting the Self-Interaction	78
11.4	Scattering in Dipole Approximation	79
12	Quantum Scattering of Light. Zero Order Approximation	81
12.1	Atom Form Factor	81
12.2	Energy flux	83
13	Light Scattering at Small Frequencies: Short-Range Scattering	84
13.1	First Order Approximation to the Ground State	84
13.2	Polarization and Dispersion: Kramers-Kronig Formula	86
13.3	Combinational Scattering	87
14	Light Scattering in Continuous Spectrum: Photoeffect	88
14.1	Radiation in Continuous Spectrum	88
14.2	Limiting Amplitude	89
14.3	Electric Current at Infinity	90
14.4	Coulomb Potential	90
14.5	Shift in Angular Distribution: Sommerfeld Formula	90
14.6	Photoeffect for Excited States	91
15	Scattering of Particles: Rutherford Formula	92
15.1	Classical Scattering by a Nucleus	92
15.2	Quantum Scattering of Electrons by the Hydrogen Atom	94
16	Hydrogen Atom in a Magnetic Field. Normal Zeemann Effect	97
16.1	Uniform Magnetic Field	97
16.2	Normal Zeemann Effect: Triplet Spectrum	98
17	Diamagnetism and Paramagnetism	100
17.1	Electric Current at Stationary States	100
17.2	Langevin Formula for Diamagnetic Susceptibility	101
17.3	Paramagnetism	101

IV	Electron Spin and Pauli Equation	103
18	Electron Spin: Experiments and Interpretation	105
18.1	Einstein-de Haas Experiment	105
18.2	Double Splitting	106
19	Pauli Equation	107
19.1	Additional Magnetic Moment	107
19.2	Additional Angular Momentum	107
19.3	Pauli Equation. Uniform Magnetic Field	108
19.4	Pauli Equation. General Maxwell Field	109
19.5	Application to the Stern-Gerlach Experiment	109
20	Einstein-de Haas and Anomalous Zeemann Effects	111
20.1	Spin-Orbital Coupling	111
20.2	Russell-Saunders Quantum Numbers	112
20.3	Landé Formula	113
20.4	Application to the Einstein-de Haas and Anomalous Zeemann Effects	114
V	Special Relativity	117
21	Electromagnetic Nature of Light	119
21.1	Maxwell Equations. Empirical Synthesis	119
21.2	Equations for Potentials	120
21.3	Problem of Luminiferous Ether: Michelson and Morley Experiment	121
21.4	Time in a Moving Frame: Lorentz transformations	122
22	The Einstein Special Relativity and Lorentz Group	123
23	Covariant Electrodynamics	126
VI	Relativistic Dirac Theory	129
24	Relativistic Equation for Electron Field	131
25	Problem of Negative Energies for Dirac Equation	133
26	Angular Momentum for Dirac Equation	134
27	Pauli Theorem	135
28	Lorentz Covariance	137
29	Lorentz Transformation of Spinors	138
29.1	Factorization of Lorentz Transformations	138
29.2	Rotations of Dirac Spinors	140
30	Coupling to Maxwell Field	142
30.1	Dirac Equation in the Maxwell Field	142
30.2	Gauge Invariance	142
31	Pauli Equation as Nonrelativistic Approximation	143
32	Charge Continuity Equation	145
33	Charged Antiparticles	146
34	Hydrogen Atom via Dirac Equation	147
34.1	Spectral Problem and Spherical Symmetry	147
34.2	Spherical Spinors and Separation of Variables	147
34.3	Radial Equations	151
34.4	Hydrogen Spectrum	152

VII	Mathematical Appendices	155
35	Newton Mechanics	157
	35.1 One Particle	157
	35.2 Many Particles	158
	35.3 Symmetry Theory	159
36	Lagrangian Mechanics	161
	36.1 One Particle	161
	36.2 Many Particles	163
37	Noether Theory of Invariants	165
	37.1 Symmetry and Noether Theorem on Invariants	165
	37.2 Application to Many-Particle Systems	166
38	Hamilton Mechanics	169
	38.1 Legendre Transform	169
	38.2 Hamilton-Jacobi Equation	170
39	Theory of Noether Currents	172
	39.1 Field Symmetry	172
	39.2 Noether Current and Continuity Equation	173
40	Application to Four Conservation Laws	176
	40.1 General Lagrangian Fields	176
	40.2 Klein-Gordon Equation	177
	40.3 Schrödinger Equation	179
41	Cauchy Problem for Maxwell Equations	181
42	Lorentz Molecular Theory of Polarization and Magnetization	184
	42.1 Constitutive Equations	184
	42.2 Stationary Molecular Fields in Dipole Approximation	184
	42.3 Non-Stationary Fields in Multipole Approximations	185
	42.4 Magnetic Moment of a Molecule	186
	42.5 Macroscopic Limit: Maxwell Equations in Matter	187
43	Long-Time Asymptotics and Scattering	189
	43.1 Retarded Potentials	189
	43.2 Limiting Amplitude Principle in Scattering Problems	190
VIII	Exercises	195
44	Exercises for Part I	197
	44.1 Exercise 1: Main Lemma of the Calculus of Variations	197
	44.2 Exercise 2: Euler-Lagrange Equations	197
	44.3 Exercise 3: Light Propagation in a Stratified Medium	198
45	Exercises for Part III	199
	45.1 Exercise 4: The Kepler Problem in 5 Easy Steps	199
	45.2 Exercise 5: Bohr-Sommerfeld Quantization of the H Atom	203
	45.3 Exercise 6: Rutherford Scattering Formula	205
	45.4 Exercise 7: Energy Flow in Maxwell Field	207
	45.5 Exercise 8: Electromagnetic Plane Waves	208
	45.6 Exercise 9: Hertzian Dipole Radiation	210
	45.7 Exercise 10: Polarizability in an External Electromagnetic Wave	212
	45.8 Exercise 11: Fresnel's Formulae for Reflection and Refraction	213
	45.9 Exercise 12: Classical Zeemann Effect	218

45.10	Exercise 13: Diamagnetism and Paramagnetism	220
45.11	Exercise 14: Stark Effect	222
45.12	Exercise 15: Vector Model for Spin-Orbital Interaction	223

0 Preface

We expose well known Schrödinger quantum mechanics with traditional applications to Hydrogen atom, but the form of the exposition is intended for a mathematically oriented reader of a graduate level.

The reader might ask why this new textbook could be useful as there are many other well established introductions to quantum mechanics. Let us explain our motivations in writing this book.

Our principal aim is to give a reasonable introduction which provides a unified mathematical strategy in applications to different problems. Many modern textbooks mainly focus on calculations of ‘matrix elements’ of a more formal nature, which are referred to as a *first approximation* of a perturbative procedure. We go one small step further in discussing the full nonperturbative problems which form the background of such calculations. This makes the strategy of the applications of the theory to specific problems more transparent.

Almost all existing texts avoid dealing with such questions, merely mentioning them. This makes the subject less accessible for understanding, and also impoverishes both its mathematical and physical aspects. On the other hand, the recognition of the status of the original nonperturbative problems leads to many questions which are open mathematical problems at the moment. Some of them are suggested by the Heisenberg Program [45]. The open discussion of these problems is also one of the principal aims of this book.

On Open Questions Quantum Mechanics exists as an axiomatic theory operating with Quantum Stationary States, Elementary Particles, Probabilities, etc. It is well established as a set of rules which give an excellent description of many experimental facts: atomic and molecular spectra, light and particle scattering, periodic system of elements, chemical reactions, etc. However, a rigorous foundation, i.e. a mathematical model of the axiomatics, is unknown at the moment because there are many open mathematical questions. The cornerstone of the theory is Schrödinger’s dynamical equation

$$(i\hbar\partial_t - e\phi(\mathbf{x}))\psi(t, \mathbf{x}) = \frac{1}{2\mu}(-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x}))^2\psi(t, \mathbf{x}) \quad (S)$$

(or, analogously, the Klein-Gordon, Dirac Eqn, etc) for the wave function $\psi(t, \mathbf{x})$, where $\phi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ are external electrostatic and magnetic vector potentials, respectively.

In a number of important particular cases, this equation could be solved exactly. Besides, the perturbation theory of the first, second, or fourth orders sometimes allows to find an approximate solution with an amazing accuracy. This gives an impression that, in principle, if one did not restrict oneself with the first order approximation, then the computations could be carried out with an arbitrarily high precision. Yet, this impression leads to a very deep confusion, since many of the fundamental quantum effects could not be described by this linear equation. Let us mention some of them.

I. Bohr’s transitions between Quantum Stationary States,

$$(0.1) \quad |E_{-}\rangle \mapsto |E_{+}\rangle,$$

II. De Broglie’s Wave-Particle Duality: diffraction of the electrons, etc.

III. Born’s probabilistic interpretation of the wave function.

The transitions (0.1) are responsible for the spectra of atom radiation which coincide with the eigenvalues of the stationary equation corresponding to (S). However the transitions are not an inherent property of the solutions of the linear equation (S).

Similarly, the equation (S) explains the diffraction pattern in the Davisson-Germer ”double-slit” experiment by the Bragg rules. However, the *discrete registration of electrons* (known as ”reduction of wave packets”), with the *counting rate* corresponding to this pattern, is not related to a property

of the solutions. Note that the stability of elementary particles is not explained yet as was pointed out by Heisenberg [45] (see [6, 35, 36] for a recent progress in this direction for nonlinear equations).

Finally, Born identified $|\psi(t, \mathbf{x})|^2$ with the density of probability just to explain the counting rate in the Davisson-Germer experiment. The identification plays the key role in quantum mechanical scattering problems. However, it has never been justified in the sense of Probability Theory.

Among other open questions: the explanation of statistic equilibrium in atom radiation (see Comment 10.1), in the photoeffect (see Comment 14.2), etc.

On Dynamical Treatment The common strategy in applications of Quantum Mechanics is a skilful combination of the postulates **I-III** with the dynamical equation (S) in the description of various quantum phenomena. The strategy is not mathematically selfconsistent since the postulates formally are not compatible with the linear autonomous equation (S). Hence, the equation requires a suitable (nonlinear) modification which would imply the postulates as inherent properties of the modified dynamics.

Fortunately, an obvious choice for the nonlinear version is well established: it is the system of coupled Maxwell-Schrödinger equations. The coupled equations are known since the Schrödinger paper [85, IV], where the charge and current densities have been expressed in terms of the wave function. Namely, the static potentials $\phi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ in (S) should be replaced by the time-dependent potentials $\phi(t, \mathbf{x})$ and $\mathbf{A}(t, \mathbf{x})$ which obey the Maxwell equations with the charge and current densities corresponding to the Schrödinger equation. The coupled equations constitute a **nonlinear system** for the electron wave function and the Maxwell field, though the Schrödinger equation is linear with respect to the wave function. The coupling is inevitable in a description of the atom spectra since the spectral lines correspond to the wavelengths of the atom electromagnetic radiation which is a solution to the Maxwell equations. Hence, the coupled equations give an authentic framework within which Quantum Mechanics, or at least some of its aspects, may be mathematically selfconsistent formulated.

We suggest a new treatment of basic quantum phenomena based on the coupled nonlinear dynamical equations. Namely, the phenomena **I** and **II** inspire the following *dynamical treatment* respectively:

A. The transitions (0.1) can be treated mathematically as the long-time asymptotics of the solutions to the coupled equations,

$$(0.2) \quad (\psi(t, \mathbf{x}), A(t, \mathbf{x})) \sim (\psi_{\pm}(\mathbf{x})e^{-i\omega_{\pm}t}, A_{\pm}(\mathbf{x})), \quad t \rightarrow \pm\infty.$$

Here $A(t, \mathbf{x}) = (\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x}))$, and the limit functions $(\psi_{\pm}(\mathbf{x})e^{-i\omega_{\pm}t}, A_{\pm}(\mathbf{x}))$ correspond to the stationary states $|E_{\pm}\rangle$. The asymptotics would mean that the set of all Quantum Stationary States is the **global point attractor** of the coupled dynamical equations (see [3, 43, 47, 95]).

B. The elementary particles seem to correspond to traveling waves (or "solitons") which are the solutions of type $(\psi(\mathbf{x} - \mathbf{v}t)e^{i\Phi(\mathbf{v}, \mathbf{x}, t)}, A(\mathbf{x} - \mathbf{v}t))$, to the coupled equations. Respectively, the de Broglie's wave-particle duality can be treated mathematically as the soliton-type asymptotics

$$(0.3) \quad (\psi(t, \mathbf{x}), A(t, \mathbf{x})) \sim \sum_k (\psi_{\pm}^k(\mathbf{x} - \mathbf{v}_{\pm}^k t)e^{i\Phi(\mathbf{v}_{\pm}^k, \mathbf{x}, t)}, A_{\pm}^k(\mathbf{x} - \mathbf{v}_{\pm}^k t)), \quad t \rightarrow \pm\infty.$$

Note that the asymptotics **A** correspond to the *bound system*, with an external confining potential, like atom with a Coulomb nuclear potential, while **B** correspond to the translation invariant systems, without an external potential. More detailed asymptotics would include also a dispersive wave which is a solution to the corresponding *free* linear system (see [64]). The dispersive wave in (0.2) should describe the electromagnetic radiation (the "photon emission") following the quantum transitions for an atom: the radiation of the wave function is impossible by the charge conservation and the neutrality of the atom.

The long time asymptotics of type **A** and **B** are proved at present time for some model nonlinear hyperbolic partial differential equations (see below). For the coupled Maxwell-Schrödinger and

Maxwell-Dirac equations, the proving of the asymptotics **A** and **B** are open problems. Only the existence of the solutions is proved for the coupled equations, [13, 42], and the existence of solitons is proved for the Maxwell-Dirac Equations, [31]. Note that the coupling of the fields is the main subject of Quantum Field Theory: Quantum Electrodynamics, etc. We consider only the *semiclassical* coupled equations and do not touch the *second quantized version* which is the subject of *Quantum Electrodynamics*. The coupled second quantized Maxwell-Dirac system of Quantum Electrodynamics is also nonlinear (see [10, 52, 82, 83]). The corresponding analogs of the conjectures are also open questions, and are very likely more difficult than in the semiclassical context.

A mathematical treatment of the Born statistical interpretation **III** is still unknown. Note that the Born identification of the counting rate with a probability is equivalent to the ergodicity of the corresponding random process. The statistical description is also necessary in the analysis of atom radiation (see Comments 10.1 and 14.2).

On Perturbative Approach The coupled equations are commonly used by perturbation techniques. The formal, perturbative approach is very successful as far as the ‘classical’ quantum mechanical results on the electron-nucleon interactions through the Maxwell field are concerned. Our above discussion mainly serves to complement this approach with a slightly more mathematical viewpoint. Our suggestion is equivalent to the explicit recognition of the status of the coupled dynamics which is responsible for the Quantum Transitions and Wave-Particle Duality. The recognition is inevitable since the both phenomena seem to be genuine nonperturbative properties of the coupled nonlinear equations.

Classical textbooks (for example, [12, 84, 89]) also use the coupled equations implicitly. That is, they treat both equations separately which corresponds to a perturbation approach for the coupled system. We follow the same strategy *explicitly* adding certain *comments* on possible relations to the coupled system and the suggested long-time asymptotics **A** and **B**. For example,

- The asymptotics **A** clarify Schrödinger’s identification of the Quantum Stationary States with eigenfunctions.
- The asymptotics **B** claim an inherent mechanism of the ”reduction of wave packets”. It clarifies the de Broglie’s wave-particle duality in the Davisson-Germer diffraction experiment and in the description of the electron beam by plane waves.

The great success of the perturbative approach to the electron-nucleon interactions, however, crucially depends on the following two main facts:

- i) the linear part of the coupled Maxwell-Schrödinger equations is completely known from the Rutherford experiment, which both detects the universal fact of the positive charge concentration and uniquely identifies the Coloumbic potential of the nuclei,
- ii) the nonlinear terms are ‘small’ due to the smallness of the Sommerfeld coupling constant $\alpha \approx 1/137$, thereby providing the numerical convergence of the perturbation series (corresponding to the *Feynman diagrams* in the *second-quantized version*).

Both of these facts no longer hold in the case of the strong nuclear interaction. Therefore, a genuine nonlinear approach might be necessary even from a more phenomenological point of view in this latter case.

On a Mathematical Justification of the Asymptotics It is natural to think that the long-time asymptotics of the type **A** and **B** are common features of a very general class of nonlinear Hamiltonian equations. Otherwise, the dynamical equations of Quantum Mechanics would be very exceptional that does not correspond to the universal character of the physical theory.

Note that the soliton-type asymptotics of type **B** have been discovered initially for *integrable equations*: KdV, sine-Gordon, cubic Schrödinger, etc. (see [77] for a survey of these results). Let us list the results on the asymptotics of type **A** and **B** obtained in the last decade for nonintegrable

nonlinear Hamiltonian equations.

The asymptotics have been proved for all finite energy solutions of i) 1D nonlinear wave equations (see [59]-[61]) for a 1D singular relativistic Klein-Gordon equation, [5]), ii) nonlinear systems of 3D wave, Klein-Gordon and Maxwell equations coupled to a classical particle (see [49]-[51] and [65]-[67]), iii) the Maxwell-Landau-Lifschitz-Gilbert Equations [56] (see [60] for a survey of these results).

For $U(1)$ -invariant 3D nonlinear Schrödinger equations, the first result on the attraction of type **A** has been established in [87, 88] (see also [79]). The first results on the asymptotics of type **B** have been established in [15, 16] for translation-invariant $U(1)$ -invariant 1D nonlinear Schrödinger equations. The results [15, 16] are extended in [18] to the dimension $n = 3$. An extension to the relativistic-invariant nonlinear Klein-Gordon equation is still an open problem (see [19, 80] for a progress in this direction). All the results [15, 16, 18, 79, 87, 88] concern initial states which are *sufficiently close to the solitary manifold*.

The *global attraction* of type **A** of *all finite energy states* is established for the first time in [63] for the $U(1)$ -invariant 1D nonlinear Klein-Gordon equation with a nonlinear interaction concentrated at one point.

In [65] an *adiabatic effective dynamics* is established, for solitons of a 3D wave equation coupled to a classical particle, in a *slowly varying external potential*. The effective dynamics explains the increment of the mass of the particle caused by its interaction with the field. The effective dynamics is extended in [35, 36] to the solitons of the nonlinear Schrödinger and Hartree equations. An extension to relativistic-invariant equations is still an open problem. On the other hand, the existence of solitons and the Einstein mass-energy identity for them are proved, respectively, in [6] and [28], for general relativistic nonlinear Klein-Gordon equations. Numerical experiments [69] demonstrate that the soliton-type asymptotics **B** hold for “all” 1D relativistic-invariant equations, however the proof is still an open problem.

There are also known results concerning stability of solitary waves for general nonlinear Hamiltonian equations, [40]. A large variety of the stability results can be found in the survey [93] concerning the nonlinear Schrödinger equations.

Note that the mathematical theory of the linear Schrödinger and Dirac equations is well established now: see for example, [7, 41, 48, 74, 81, 96]. However, the asymptotics **A** and **B** generally do not hold for the linear autonomous Schrödinger and Dirac equations.

Main Goals of the Exposition We pursue the following two principal goals:

I. To explain why the theory has admitted its present form of a dynamical system described by the Schrödinger, Pauli or Dirac equations coupled to the Maxwell equations.

We follow all details of the development of the coupled dynamical equations, from experimental facts to the related mathematical context. We introduce the Schrödinger equation as a wave equation for which rays coincide with trajectories of the Lorenz equation for the classical electron. This introduction is based on the *geometric optics* and WKB short wave asymptotics, and is close to the original Schrödinger’s idea on the Hamilton *optical-mechanical analogy*, [85, II]. We explain in detail the geometric optics and the WKB asymptotics. For the introduction of the Pauli equation we analyze the double splitting in the Stern-Gerlach experiment, the Einstein-de Haas effect and the anomalous Zeemann effect. The Dirac equation corresponds to a relativistic energy-momentum relation similarly to the Schrödinger equation which corresponds to a non-relativistic one.

II. To demonstrate that the coupled dynamical equations allow us to describe the basic quantum phenomena of interaction of matter and electromagnetic radiation as inherent properties of the dynamics if the asymptotics **A** and **B** would hold.

Our analysis shows that the asymptotics would play the key role in a mathematical foundation of Quantum Mechanics. Our exposition cannot be rigorous when we solve the nonlinear coupled Maxwell-

Schrödinger equations in the first order approximation. We try to be careful in the solution of the corresponding linear problems but we did not strain to be everywhere mathematically rigorous in order not to overburden the exposition.

Among Other Novelties of the Exposition are the following:

- An application of the Lagrangian formalism for the identification i) of the energy, momentum and angular momentum for the Schrödinger equation (Lecture 6), ii) of the coupling of the Maxwell and Schrödinger equations (Lecture 7, cf. [12, 84, 101]).
- The straightforward derivation of the *Rydberg-Ritz combination principle* and *intensities for the dipole radiation* from the coupled Maxwell-Schrödinger equations (Lecture 10). The derivation is a formalized version of the approach [89].
- An update version of the Lorentz theory of molecular polarization and magnetization (Lecture 42), necessary for the quantum theory of dispersion and diamagnetism. We follow mainly [4] applying the theory of distributions. The theory gives a framework for quantum versions of the Kramers-Kronig dispersion theory (Lecture 14) and the Langevin theory of diamagnetism (Lecture 18).
- An update explanation of the *red bound* for the *photoeffect* by the scattering in the continuous spectrum. We calculate the Wentzel formula for the photoeffect as the *limiting amplitude* of the scattering via the *limiting absorption principle* (Lecture 15, cf. [89]). Then the *Einstein formula* for the energy of the photoelectrons follows from the *long range asymptotics* of the limiting amplitude.
- A concise presentation of the Russell-Saunders theory of the coupling between the orbital and spin angular momentum (Lecture 21). We follow mainly [8, 17, 71, 84] and comment on probable relations with the coupled Maxwell-Pauli equations.
- An update form of the Noether Theorem on Currents with the complete proof (Lecture 40, cf. [37, 103]).

We demonstrate a parallelism of quantum and classical description to clarify their relations and to motivate an introduction of the corresponding quantum phenomenology: for example, the classical and quantum description of diamagnetism, Zeemann effect, scattering of light and particles, the introduction of *differential cross section*, *magnetic moment*, etc.

We explain carefully all details of the calculations and all necessary methods of modern Mathematical Physics: the Lagrange and Hamilton theory for the fields, the Maxwell Electrodynamics and the Einstein Special Relativity, scattering theory and representation theory of the rotation group. Let us note that we consider only one-electron problems (Hydrogen atom, alkali atoms, etc). We do not touch multi-particle problems, Hartree-Fock methods, etc (see [17, 21, 72, 94]).

Further Reading Our main goal is a concise explanation of basic theoretical concepts. More technical details and a systematic comparison with the experimental data can be found in [9, 17, 84, 89].

Plan of Exposition In the Introduction we describe the chronology of the conceptual development of Quantum Mechanics. Then in Lectures 2-4 we provide the mathematical background for a concise introduction of the Schrödinger equation in Lectures 5-7. Lectures 8-21 concern various applications of the Schrödinger and Pauli equations. The relativistic Dirac theory is exposed in Lectures 22-35, where we benefit a lot from the book of Hannabus [44]. In Lectures 36-43 we collect mathematical appendices. In Sections 44-45 we solve numerous exercises containing main achievements of the “old quantum mechanics”. Let us explain the plan of the lectures in some detail.

Quantum Chronology The genesis of the Schrödinger equation has been inspired by the lack of a matter equation in classical electrodynamics. The Schrödinger theory is the result of a synthesis of a theoretical development with various experimental observations. The “quantum chronology” starts from Kirchhoff’s *spectral law* (1859) and the invention of the *vacuum tube* by Crookes (1870). The next main steps are the identification of the *cathode rays* in the vacuum tube with the electrons by

Thomson (1897), and the fundamental Planck relation $E = \hbar\omega$ (1901) inspired by the comparison of the experimental Wien formula for the spectral density in the Kirchhoff law, with the Boltzmann-Gibbs canonical distribution. The Planck relation has been applied by Einstein to the photoeffect (1905). The Einstein theory has been developed further by Bohr (1913) to explain the Rydberg-Ritz combination principle and the stability of Rutherford's planetary model for the atom. Finally, Bohr's theory of *atom stationary states* has led to the *quantum rules* of Debye and Sommerfeld-Wilson (1913 and 1916) for the *action function*, which is a solution to the Hamilton-Jacobi equation. The quantum rules provide the atom stationary states and allow to reproduce some basic experimentally observed spectral lines.

In 1926 Schrödinger introduced an equation for the wave function [85]. The equation extends de Broglie's wave-particle theory (1922) from free particles to bound particles. De Broglie's theory is based on an algebraic argument relying on Einstein's special relativity and the Planck relation. Schrödinger's extension combines the algebraic argument with Hamilton's optical-mechanical analogy. The analogy is justified by the WKB short wave asymptotics for the solutions to the Schrödinger equation. That is, the corresponding "eikonal equation" coincides with the Hamilton-Jacobi equation of the Maxwell-Lorentz theory for the classical electron in the Maxwell field. This means that the Schrödinger equation describes the cathode rays as short-wave solutions. An important role of the Hamilton-Jacobi equation has been recognized previously related to the Debye-Sommerfeld-Wilson quantization rules for the action.

Formally, a wave equation of the Schrödinger type can be introduced directly by an identification of the trajectories of the classical electron in the Maxwell field with the rays of short-wave solutions. However, the magnitude \hbar of the corresponding small parameter, and the central role of Quantum Stationary States, can be recognized only from the whole development, starting with the Kirchhoff spectral law and leading to de Broglie's wave-particle theory.

Mathematical Background In Lectures 2-4 we expose the Lagrangian field theory, and the Lagrangian form of Maxwell's electrodynamics. Lecture 5 concerns the *geometric optics* and short-wave WKB asymptotics. In Lecture 6 we introduce the Schrödinger equation, the Heisenberg representation, and prove conservation laws. In Lecture 7 we derive the coupling of the Schrödinger wave function to the Maxwell field through the Lagrangian formalism.

Applications In Lectures 8-18 we apply the Schrödinger-Maxwell equations to the derivation of the spectrum of the hydrogen atom, the *dipole radiation* and *selection rules*, the *differential cross sections* of the scattering of light and particles by an atom, the *refraction coefficient*, the *diamagnetism*, and the *normal Zeemann effect*. Lecture 9 contains a complete theory of *quantum angular momentum*, including the *representations* of the corresponding *Lie algebra*. In Lecture 15 we explain the *photoeffect* by the *limiting amplitude principle* for the scattering in the *continuous spectrum*. In Lectures 19-21 we introduce the *Pauli equation* with *electron spin* and apply it to the calculation of the *gyromagnetic ratio*. In Lectures 22-35 we expose the *special relativity* (Lorentz Transformations and Covariant Electrodynamics), introduce the relativistic *Dirac equation*, prove its covariance and the *Pauli theorem*, and calculate the spectrum of the hydrogen atom via the Dirac equation.

Mathematical Appendices In Lectures 36-43 we give an introduction to the Lagrange theory for finite-dimensional systems and for fields: variational principle, conservation of energy, momentum, angular momentum, Noether invariants, Hamilton equations, Hamilton-Jacobi theorem. In Lecture 40 we give a new simple proof of the Noether theorem on currents. In Lecture 42 we expose the Lorentz theory of molecular polarization and magnetization. In Lecture 43 we explain the limiting amplitude principle, limiting absorption principle and the role of retarded potentials for the calculation of limiting amplitudes.

Exercises In Part VIII, written by C.Adam, we collect the solutions to the related classical prob-

lems: the Kepler problem, Bohr-Sommerfeld quantization, energy and momentum in the Maxwell field, electromagnetic plane waves and Fresnel's formulae, Hertzian dipole radiation, vector model for the spin-orbital coupling, etc.

Acknowledgments I thank H.Brezis, I.M.Gelfand, P.Joly, J.Lebowitz, E.Lieb, I.M.Sigal, A.Shnirelman, M.Shubin, H.Spohn, M.I.Vishik and E.Zeidler for fruitful discussions. I thank very much Dr. C.Adam for his help in preparation of this book. I am indebted to Mechanical-Mathematical Department of Moscow University for the support of my research. Finally, I thank the Faculty of Mathematics of Vienna University, the Max-Planck Institute for Mathematics in the Sciences (Leipzig), the Project ONDES (INRIA, Rocquencourt), and Wolfgang Pauli Institute for their hospitality.

A.Komech

Moscow, 12.10.2005

1 Introduction: Quantum Chronology 1859-1927

1.1 Missing “Matter Equation”

We mark important points of the development of Quantum Theory. Every point is significant in either obtaining new experimental results or inventing a new treatment, or developing new mathematical methods.

The Maxwell theory (1865) perfectly describes the motion of charged particles in a *given* electromagnetic field and also the propagation of electromagnetic waves generated by *known* charge and current densities ρ, j . However, generally it cannot describe the **simultaneous** evolution of the unknown densities and fields since the *microscopic* evolution equation for the densities is missing. The representation of the densities as an aggregation of moving charged particles does not help, since the corresponding *mass density* μ is unknown. That is, the ratio ρ/μ is not known and even does not have a reasonable meaning, since the ratio e/m takes different values for different elementary particles.

The situation is better on the *macroscopic* level in *simple* media with known electric and magnetic permeability and conductivity, since then the the *macroscopic* charge and current densities $\rho_{\text{mac}}, j_{\text{mac}}$ are *functions* of the fields (*polarization, magnetization, Ohm’s law*, etc). However, this is not the case for the vacuum. Hence, classical electrodynamics is not sufficient to explain the structure of matter at the *microscopic* level.

Quantum mechanics just provides various matter equations: Schrödinger, Klein-Gordon, Dirac equations, etc. The equations arise *inside* classical electrodynamics, thermodynamics, optics and atomic physics from experimental observations of various aspects of the field-matter interaction and their theoretical treatment. Let us briefly sketch the chronology of the development.

1.2 Thermodynamics, Optics and Electrodynamics

Kirchhoff-Planck: The Black-Body Radiation

In 1859 Kirchhoff stated the radiation law which predicts a specific spectral density of light waves radiated by a black body at a fixed temperature. Light was identified with the electromagnetic field by Maxwell in 1865. Hence, the Kirchhoff law concerns the spectrum of the equilibrium distribution of the electromagnetic field at a fixed temperature. Therefore, it provides a deep indirect information on the interaction of matter with the Maxwell field.

The experimental measurements have been performed by Tyndall in 1865, Crova in 1880, Langley in 1886, Weber in 1887, and Paschen in 1895-1899. Very precise measurements were made in 1899 by Lummer and Pringsheim, and Kurlbaum and Rubens. They confirmed the Wien formula (1896)

$$(1.4) \quad I(\omega) \sim \omega^3 \exp\left(-\frac{\beta\omega}{T}\right).$$

Note that the traditional reference to the *black* body just means that its equilibrium radiation coincides with the equilibrium Maxwell field since the absorption of the black body is zero by definition. The comparison of (1.4) with the general equilibrium Boltzmann-Gibbs distribution $\exp\left(-\frac{E}{kT}\right)$ (where k is the Boltzmann constant) suggests the famous Planck relation (1901)

$$E = \hbar\omega, \quad (P)$$

where E is the energy of the “emitted photon” and $\hbar = k\beta \approx 1.05 \cdot 10^{-27} \text{erg} \cdot \text{sec}$ is the Planck

constant. Using this relation, Planck adjusted the formula (1.4) as

$$I(\omega) \sim \omega^2 \frac{\exp(-\frac{a\omega}{T})}{1 - \exp(-\frac{a\omega}{T})}. \quad (KP)$$

Rydberg-Ritz: Atom Spectra

Atom spectra provide extremely important information on the structure of the atom. In 1885 Balmer discovered the representation $\omega_{2n} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$ ($n \geq 3$) for a spectral series in the spectrum of the hydrogen atom. Later, similar representations were found for other series by Paschen (1908) $\omega_{3n} = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right)$ ($n \geq 4$), Lyman (1909) $\omega_{1n} = R\left(1 - \frac{1}{n^2}\right)$ ($n \geq 2$), and Brackett (1914) $\omega_{4n} = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right)$ ($n \geq 5$). Similar structure

$$\omega_{mn} = \omega_m - \omega_n, \quad (R)$$

has been discovered experimentally by Rydberg (1900) for all the lines in several series of other elements. The importance of these observations was also stressed by Ritz (1908), so it is now commonly known as the *Rydberg-Ritz combination principle*, and the numbers ω_m are called *terms*.

Crookes-Herz-Perrin-Thomson: The Cathode Rays and the Electron

The cathode rays were discovered first in *vacuum tube* by Crookes in 1870 (he made a *discharge tube* with a vacuum level higher than that of the *Geissler tube* used by Faraday in 1836-1838). The rays demonstrated the continuous motion of charge in the vacuum in the presence of a Maxwell field. This is just one of the situations which is not covered by classical electrodynamics.

The deflection of cathode rays in a magnetic field has been observed in 1880-1890 by Hertz, Lenard, Perrin and many others. Some physicists thought, like Goldstein, Hertz, and Lenard, that this phenomenon is like light, due to vibrations of the ether or even that it is light of short wavelength. It is easily understood that such rays may have a rectilinear path, excite phosphorescence, and effect photographic plates. Others thought, like Crookes, J.J. Thomson, Perrin and others, that these rays are formed by matter which is negatively charged and moving with great velocity, and on this hypothesis their mechanical properties, as well as the manner in which they become curved in a magnetic field, are readily explicable. In 1895, Perrin collected the cathode rays, obtaining a negative charge.

In 1897 J.J. Thomson showed that the rays are also deflected by an electrostatic field. He systematized all previous observations and demonstrated the particle-like behavior of the cathode rays which is described by the Lorentz equation,

$$a = \frac{e}{\mu}(E + v \times B), \quad (L)$$

where $\frac{e}{\mu} < 0$. Concretely, he identified the cathode rays with a beam of particles with negative charge and introduced the name *electron* for these particles. This study led to the first measurement of the ratio $\frac{e}{\mu}$ close to its present value. Kauffmann [58] also observed the magnetic deflection of cathode rays and obtained a ratio $\frac{e}{\mu}$ which is close to the value of J.J. Thomson.

J.J. Thomson's identification of the cathode rays led to many fundamental problems concerning the size and the structure of the electron. In particular,

i) Abraham (1906) noted that the energy and the mass of the electron are infinite if its radius is zero. He introduced the model of the *extended* electron and calculated its radius.

ii) The extended electron cannot be stable because of the electrostatic repulsion (Poincaré 1908).

So, classical electrodynamics had to be complemented with a *matter equation* which could describe i) the cathode rays and their *particle-like* behavior, and ii) the stability of the electron.

Herz-Einstein: Photoeffect

In 1887 Herz discovered the photoeffect (the "light electricity") via the generation of electric charge by the sun radiation. Later, the photoeffect has been observed with different types of electromagnetic radiation by Stoletov, Elster, Geitel, Righi, Townsend, Rutherford, Compton and many others. The experimental observations led to the relation

$$\hbar\omega = E_{\text{el}} - A. \quad (E)$$

Here, E_{el} is the (maximal) energy of the photoelectrons detached from the metal by light of the frequency ω . The constant A depends on the metal. The photoeffect occurs only for large frequencies of light $\omega > \omega_{\text{red}} = A/\hbar$, where ω_{red} is called the "red bound" of the photoeffect.

In 1905 Einstein proposed the theory of the photoeffect [30]: he identified the relation (E) with energy conservation. That is, Einstein

I. Identified the quantity $\hbar\omega$ with the energy of the *absorbed* photon with frequency ω in accordance with the Planck relation (P) (which concerns the *emitted* photon !), and

II. Identified A with the *escape energy* of the metal.

This explanation treats light as a collection of particle-like "photons" that cannot be explained by using a wave picture of light and the classical representation of electrons as particles.

1.3 Atomic Physics

Rutherford: The Nucleus of the Atom and Atom Stability

In 1913 Rutherford discovered the nucleus of the atom in an experiment on the scattering of α -particles. This discovery suggested to him the classical model of the atom, where a finite number of electrons moves around a point-like nucleus with positive charge. The electrons are governed by the classical Lorentz eq. (L). However, the model is unstable due to the radiation of the rotating electrons in accordance with Maxwell electrodynamics. Therefore, the Maxwell theory is insufficient to explain the stability of the atoms.

Bohr: Quantum Stationary States and Transitions

In 1913 Niels Bohr has proposed a new phenomenology for description of the atom stability. Namely, he represented the Rydberg-Ritz combination principle (R) in the form

$$(1.5) \quad \hbar\omega_{mn} = E_m - E_n,$$

which was suggested by the comparison of (R) with the Planck relation (P) and the Einstein treatment of the relation for the photoeffect, (E). Moreover, Bohr interpreted (B) generalizing the Planck and Einstein ideas:

I. For an atom, there exist **Stationary States** $|E_n\rangle$ with the energies E_n . The atom is "always" in a stationary state, and, sometimes, make **transitions** (or "jumps") from one Stationary State to another,

$$(1.6) \quad |E_m\rangle \mapsto |E_n\rangle.$$

II. The transition is accompanied by the radiation or absorption of light with frequency ω_{mn} .

III. The identity (1.5) is energy balance in the transition, in accordance with the identification of Planck and Einstein of the quantum $\hbar\omega_{mn}$ with the energy of an emitted or absorbed photon.

Both, the role of the Planck constant \hbar and the discreteness of the energies E_n of the stationary states, cannot be explained by the Maxwell theory. The discreteness of the energies is related to a restriction to certain *stable* orbits of the electron in the atom.

Debye-Sommerfeld-Wilson: “Old Quantum Theory”

In 1913 Debye stated the *quantum rule* for the determination of stable *periodic* orbits of the electrons in the atom,

$$\Delta S = 2\pi n\hbar, \quad n = 1, 2, 3, \dots \quad (D)$$

where ΔS is the action integral corresponding to the time-periodic orbit of the electron. The rule was motivated by the Ehrenfest idea of adiabatic invariance. The quantum rules allowed to find the hydrogen spectral terms $\omega_n = \frac{R}{n^2}$, $n = 1, 2, \dots$ which exactly agree with the series of Lyman, Balmer etc. In 1916 Sommerfeld and Wilson extended the rule to more general *quasiperiodic* orbits.

In 1923 Bohr has developed the *correspondence principle* which allowed him to discover the *selection rules* for the *magnetic and azimuthal quantum numbers*. The selection rules play the key role in the explanation of atom spectra and agree with experimental observations, see [12] and [89, Vol.I].

Zeemann-Stern-Gerlach: Atoms in Magnetic Fields

In 1896 Zeemann discovered the splitting of the spectral lines of atoms in a magnetic field. Lorentz explained the splitting by the Maxwell theory in the simplest case of the *normal Zeemann effect* when the line ω splits into three lines: ω and $\omega_{\pm} = \omega \pm \Delta\omega$, where $\Delta\omega$ is proportional to the magnetic field. However, the explanation of the general *anomalous Zeemann effect* cannot be deduced from the Maxwell theory. One example is the double splitting of the spectra of alkali atoms.

The Maxwell theory predicts a unique value r for the gyromagnetic ratio $|\mathbf{m}|/|\mathbf{J}| = \mu_B$, where $\mu_B = \frac{|e|\hbar}{2mc}$ is the *Bohr magneton* and $|\mathbf{m}|$ resp. $|\mathbf{J}|$ are the magnetic and mechanical momenta of the electron in an atom. In 1915 Einstein and de Haas first measured the gyromagnetic ratio by an observation of a magnetization of an iron in an external weak magnetic field. However, the observed ratio was $2r$, i.e., two times larger than the theoretical value.

In 1921 Stern and Gerlach observed the double splitting of a beam of silver atoms in a strong non-uniform magnetic field. This implies that the stationary state of the atom is split into two states with different gyromagnetic ratios $|\mathbf{m}|/|\mathbf{J}|$, which again contradicts the Maxwell theory.

Compton: Scattering of Light by Electrons

In 1923, Compton discovered that the scattered light has a wavelength λ' different from the wavelength λ of the incident light:

$$\lambda' - \lambda \sim \frac{2\hbar}{\mu c} \sin^2 \frac{\theta}{2},$$

where θ is the angle between the incident and scattered waves, and μ is the electron mass. Similarly to the photoeffect, the scattering also cannot be explained by using a wave picture of light, where the wavelength does not change.

De Broglie: Wave-Particle Duality for Free Particles

In 1924 de Broglie, in his PhD thesis, introduced a wave function for a possible description of matter by waves, [55], in analogy with the particle-wave duality of light which is demonstrated by the Maxwell theory and the photoeffect. Namely, he has applied Einstein's Special Relativity Theory to a beam of free particles with the energy-momentum vector (E, \mathbf{p}) .

I. The beam is identified with a plane wave by the following “wave-particle” relation:

$$(1.7) \quad \psi(t, \mathbf{x}) = C e^{i(\mathbf{k}\mathbf{x} - \omega t)} \leftrightarrow \text{beam of free particles.}$$

II. The Einstein relativity principle and the Planck relation (P) imply the identity

$$(1.8) \quad (E, \mathbf{p}) = \hbar(\omega, \mathbf{k}).$$

The identity plays a crucial role everywhere in quantum theory. In particular, it implies the famous de Broglie relation for the “particle wave length” $\lambda = 2\pi/|\mathbf{k}|$,

$$\lambda = \frac{2\pi\hbar}{|\mathbf{p}|}.$$

It also implies the relativistic dispersion relation

$$(1.9) \quad \frac{\hbar^2 \omega^2}{c^2} = \hbar^2 \mathbf{k}^2 + \mu^2 c^2,$$

where μ is the particle mass and c is the speed of light. It follows from the expression for the Hamiltonian of the relativistic particle (see (38.6))

$$(1.10) \quad \frac{E^2}{c^2} = \mathbf{p}^2 + \mu^2 c^2.$$

iv) For small values of $|\mathbf{p}| \ll \mu c$ the non-relativistic approximation holds,

$$(1.11) \quad E = \sqrt{\mathbf{p}^2 c^2 + \mu^2 c^4} \approx \mu c^2 + \frac{\mathbf{p}^2}{2\mu}.$$

Dropping here the “unessential” additive constant μc^2 , we get the non-relativistic dispersion relation

$$(1.12) \quad \hbar\omega = \frac{\hbar^2 \mathbf{k}^2}{2\mu}.$$

The dispersion relations (1.9) resp. (1.12) implies the *free* Klein-Gordon resp. Schrödinger equation for the corresponding wave function $\psi(t, \mathbf{x}) = e^{i(\mathbf{k}\mathbf{x} - \omega t)}$:

$$\frac{1}{c^2} [i\hbar\partial_t]^2 \psi(t, \mathbf{x}) = [(-i\hbar\nabla_{\mathbf{x}})^2 + \mu^2 c^2] \psi(t, \mathbf{x}), \quad (KG_0)$$

$$i\hbar\partial_t \psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar\nabla_{\mathbf{x}}]^2 \psi(t, \mathbf{x}). \quad (S_0)$$

Klein-Gordon-Schrödinger: Wave Equation for Bound Particles

In 1925-1926 Klein, Gordon and Schrödinger extended de Broglie's wave equation to the bound electron in an external Maxwell field. The *free* equations (KG_0) resp. (S_0) formally follow from the energy-momentum relations (1.10) resp. (1.11) by the substitutions

$$(1.13) \quad E \mapsto i\hbar\partial_t, \quad \mathbf{p} \mapsto -i\hbar\nabla_{\mathbf{x}}.$$

For an electron in the external scalar potential $\phi(t, \mathbf{x})$ and magnetic vector potential $\mathbf{A}(t, \mathbf{x})$, the (conserved) energy E is given by $[E - e\phi(t, \mathbf{x})]^2/c^2 = [\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2 + \mu^2c^2$, where e is the charge of the electron (see (4.36)). Then Klein, Gordon and Schrödinger generalized (KG_0) to

$$\frac{1}{c^2}[i\hbar\partial_t - e\phi(t, \mathbf{x})]^2\psi(t, \mathbf{x}) = [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi(t, \mathbf{x}) + \mu^2c^2\psi(t, \mathbf{x}). \quad (KG)$$

Schrödinger also generalized the nonrelativistic approximation (1.12) to

$$E - e\phi(t, \mathbf{x}) = (\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x}))^2/(2\mu),$$

which transforms into the wave equation

$$[i\hbar\partial_t - e\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi(t, \mathbf{x}). \quad (S)$$

The next crucial step of Schrödinger's theory is the identification of stationary states with solutions of the type $\exp(-i\omega t)\psi(\mathbf{x})$ for the static external Maxwell fields $\phi(t, \mathbf{x}) \equiv \phi(\mathbf{x})$ and $A(t, \mathbf{x}) \equiv \mathbf{A}(\mathbf{x})$. This identification is suggested by the de Broglie plain wave $\exp(-i\omega t)\exp(i\mathbf{k}\mathbf{x})$, where only the spatial factor has to be modified since the external field "twists" space but not time. The energy is again $E = \hbar\omega$. This identification leads to the corresponding *stationary equations* which are the *eigenvalue problems*,

$$\begin{aligned} \frac{1}{c^2}[\omega - e\phi(\mathbf{x})]^2\psi(\mathbf{x}) &= [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]^2\psi(\mathbf{x}) + \mu^2c^2\psi(\mathbf{x}), \\ [\omega - e\phi(\mathbf{x})]\psi(\mathbf{x}) &= \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]^2\psi(\mathbf{x}) \end{aligned}$$

for the determination of the energies $E = \hbar\omega$ and the amplitudes $\psi(\mathbf{x})$ of the stationary states.

Schrödinger calculated all solutions to the last equation for the hydrogen atom: $\phi(\mathbf{x}) = -e/|\mathbf{x}|$ is the Coulomb potential of the nucleus, and $\mathbf{A}(\mathbf{x}) = 0$. The agreement with the experimentally observed spectrum was perfect. The calculation uses the standard separation of variables in spherical coordinates, which involves some integer numbers as in the Debye quantum condition (D). It was just this analogy which suggested to Schrödinger an eigenvalue problem for the determination of the stationary states of the atom.

Heisenberg: Matrix Mechanics

In 1925 Heisenberg have extended significantly the Bohr correspondence principle. He assigned the *infinite matrix* i.e. the operator in the Hilbert space, to each classical observable: energy, momentum, coordinate, angular momentum, etc. This assignment is called now *the quantization* of the corresponding classical dynamical system. In particular, the substitutions (1.13) give an example of the assignment at a fixed time. In this *matrix mechanics* the operators depend on time and obeys the dynamical equations which formally coincide with the equations for the classical observables. The

Heisenberg approach allowed him to discover the key **Uncertainty Principle** (see Lecture 7) which plays a fundamental role in the theoretical analysis and applications of the quantum theory.

The Heisenberg ideas were developed by Born, Jordan and others. Later on it was proved that the Heisenberg approach is completely equivalent to the Schrödinger theory. The equivalence has played a crucial role in further development of the quantum theory leading to the *second quantization* and *quantum field theory*.

Uhlenbeck-Goudsmith-Pauli: Nonrelativistic Theory of Spin

In 1925 Uhlenbeck and Goudsmith introduced the hypothesis of the existence of the *spin* of the electron, i.e., of its *own angular momentum* \mathbf{s} (i.e. not related to its rotation) and a magnetic moment \mathbf{m} (not related to the corresponding convection current) with the gyromagnetic ratio $g := |\mathbf{m}|/|\mathbf{s}| = 2\mu_B$. The hypothesis had been inspired by the double splitting in the Stern-Gerlach experiment, the *anomalous Zeemann effect* and the Einstein-de Haas experiment. In 1927 Pauli obtained the wave equation which takes the spin of the electron into account,

$$(1.14) \quad [i\hbar\partial_t - e\phi(t, \mathbf{x})]\Psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\Psi(t, \mathbf{x}) + g\mu_B \sum_1^3 \hat{\mathbf{s}}_k \mathbf{B}_k \Psi(t, \mathbf{x}),$$

where the wave function $\Psi(t, \mathbf{x}) = (\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x}))$ with two complex-valued functions $\psi_j(t, \mathbf{x})$. Further, \mathbf{B}_k are the components of the *uniform* magnetic field and $\hat{\mathbf{s}}_k = \sigma_k/2$, where σ_k are the complex 2×2 *Pauli matrices*. Pauli obtained the equation by just postulating the double splitting, the *gyromagnetic ratio* $g = 2$, and covariance with respect to space rotations.

This equation leads to the correct gyromagnetic ratio observed in the Einstein-de Haas experiment. It also explains the anomalous Zeemann effect and the Stern-Gerlach double splitting. The agreement with many experimental observations was a great triumph of the quantum theory.

Dirac: Relativistic Theory of Spin

In 1927 Dirac discovered the relativistic invariant equation which generalizes the Pauli and Klein-Gordon equations,

$$\sum_{\alpha=0}^3 \gamma_{\alpha} [i\hbar\nabla_{\alpha} - \frac{e}{c}\mathcal{A}_{\alpha}(x)] = \mu c \psi(x), \quad x \in \mathbb{R}^4,$$

where $\nabla_0 = \partial_t/c$, $(\nabla_1, \nabla_2, \nabla_3) = \nabla_{\mathbf{x}}$, $\mathcal{A}_0(x) = \phi(x)$ and $\mathcal{A}_k = -\mathbf{A}_k(x)$, $k = 1, 2, 3$, γ_{α} are the 4×4 Dirac matrices, and $\psi(x) \in \mathbb{C}^4$ for $x \in \mathbb{R}^4$.

The Dirac equation automatically provides the correct gyromagnetic ratio $g = 2$ for the electron. It gives a much more precise description of the hydrogen atom spectrum than the Schrödinger and Pauli equations.

Davisson-Germer: Interference of Electrons

In 1927 Davisson and Germer observed the interference of electron beams. Later the experiments were repeated and confirmed by many authors: Thomson, Rupp, Kikouchi and others. In 1949 Biberman, Sushkin and Fabrikant observed the interference pattern with a weak beam with a very low rate of registration of the electrons.

Born: The Probabilistic Interpretation of the Wave Function

In 1927 Born proposed the following interpretation of the wave function to explain the Davisson-Germer experiment: $|\psi(t, \mathbf{x})|^2$ is a *density of probability*.

Part I

Lagrangian Field Theory

2 Euler-Lagrange Field Equations

We introduce two equations important for Quantum Mechanics, namely, the Klein-Gordon and Schrödinger equations. We prove that they are of Lagrangian form. Then we introduce an action functional and prove the Hamilton least action principle.

The quantum mechanical evolution equations for charged matter in an electromagnetic field are provided by the Schrödinger equation and Klein-Gordon equation in the non-relativistic and relativistic cases, respectively. Both equations may be derived from an action functional via the Hamiltonian least action principle, and, therefore, a Lagrangian description exists for both cases.

2.1 Klein-Gordon and Schrödinger Equations

Let us define $x_0 = ct$, $\mathbf{x} = (x_1, x_2, x_3)$, $x = (x_0, \mathbf{x})$ and consider the Klein-Gordon equation

$$(2.1) \quad \begin{aligned} & [i\hbar\nabla_0 - \frac{e}{c}\phi(x)]^2\psi(x) \\ & = [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(x)]^2\psi(x) + \mu^2c^2\psi(x), \quad x \in \mathbb{R}^4, \end{aligned}$$

where the function $\psi(x)$ takes complex values and $\mu > 0$ is the electron mass. Let us also consider the Schrödinger equation

$$(2.2) \quad \begin{aligned} & [i\hbar\nabla_0 - e\phi(x)]\psi(x) \\ & = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(x)]^2\psi(x), \quad x \in \mathbb{R}^4, \end{aligned}$$

where $x_0 := t$.

2.2 Lagrangian Density

Definition 2.1 We will identify the complex vectors $\psi \in \mathbb{C}^M$ with the real vectors $\mathbb{R}\psi := (\operatorname{Re}\psi, \operatorname{Im}\psi) \in \mathbb{R}^{2M}$ and the multiplication by a complex number with an application of the corresponding matrix. We will denote by \cdot the real scalar product in \mathbb{R}^{2M} .

This definition implies the formulas

$$(2.3) \quad \mathbb{R}u \cdot \mathbb{R}v = \operatorname{Re}(u\bar{v}),$$

$$(2.4) \quad \nabla_{\mathbb{R}u}(u \cdot iv) = iv, \quad \nabla_{\mathbb{R}v}(u \cdot iv) = -iu, \quad \nabla_{\mathbb{R}u}(iu \cdot v) = -iv, \quad \nabla_{\mathbb{R}v}(iu \cdot v) = iu,$$

for $u, v \in \mathbb{C}$ since $u \cdot iv = -iu \cdot v$ and $iu \cdot v = -u \cdot iv$.

Let us introduce the *Lagrangian densities* \mathcal{L} for Eqs. (2.1) and (2.2) as the *real* functions defined by (cf. (35.3), (35.10)),

$$(2.5) \quad \begin{aligned} \mathcal{L}_{KG}(x, \psi, \nabla\psi) &= \frac{|[i\hbar\nabla_0 - \frac{e}{c}\phi(x)]\psi|^2}{2} - \frac{|[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)]\psi|^2}{2} \\ &\quad - \mu^2c^2\frac{|\psi|^2}{2}, \end{aligned}$$

$$(2.6) \quad \mathcal{L}_S(x, \psi, \nabla\psi) = [i\hbar\nabla_0 - e\phi(x)]\psi \cdot \psi - \frac{1}{2\mu}|[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)]\psi|^2.$$

We will demonstrate below that the field equations (2.1), (2.2) can be represented in the **Euler-Lagrange** form,

$$(2.7) \quad \mathcal{L}_\psi(x, \psi(x), \nabla\psi(x)) - \sum_{\alpha=0}^3 \nabla_\alpha \mathcal{L}_{\nabla_\alpha \psi}(x, \psi(x), \nabla\psi(x)) = 0, \quad x \in \mathbb{R}^4,$$

where \mathcal{L} is the corresponding Lagrangian density.

Remark 2.2 In (2.5) – (2.7) the x and $\psi, \nabla_\alpha \psi$ are considered as independent variables with the values in \mathbb{R}^4 and \mathbb{R}^2 , respectively.

Definition 2.3 The Lagrangian field $\psi(x)$ with values in \mathbb{R}^N is the dynamical system described by the N real scalar equations (2.7) with a given Lagrangian density $\mathcal{L}(x, \psi, \nabla\psi) \in C^2(\mathbb{R}^4 \times \mathbb{R}^N \times \mathbb{R}^{4N})$.

Definition 2.4 The **canonically conjugate fields** $\pi_\alpha(x)$ are defined by

$$(2.8) \quad \pi_\alpha(x) = \mathcal{L}_{\nabla_\alpha \psi}(x, \psi(x), \nabla\psi(x)), \quad \alpha = 0, \dots, 3.$$

With these notations the Euler-Lagrange equations (2.7) read

$$(2.9) \quad \nabla_\alpha \pi_\alpha(x) = \mathcal{L}_\psi(x, \psi(x), \nabla\psi(x)), \quad x \in \mathbb{R}^4.$$

Here and below we use the Einstein convention $\nabla_\alpha \pi_\alpha(x) := \sum_\alpha \nabla_\alpha \pi_\alpha(x)$ etc. Also $\alpha = 0, 1, 2, 3$ and $k = 1, 2, 3$.

2.3 Free Equations

First consider the *free* equations without Maxwell field and without nonlinear self-interaction, and with $\hbar = 1$:

$$(2.10) \quad \nabla_0^2 \psi(x) = \nabla_{\mathbf{x}}^2 \psi(x) - \mu^2 c^2 \psi(x), \quad x \in \mathbb{R}^4,$$

$$(2.11) \quad -i \nabla_0 \psi(x) = \frac{1}{2\mu} \nabla_{\mathbf{x}}^2 \psi(x), \quad x \in \mathbb{R}^4,$$

Then the Lagrangian densities (2.5), (2.6) become

$$(2.12) \quad \mathcal{L}_{KG}^0(x, \psi, \nabla\psi) = \frac{|\nabla_0 \psi|^2}{2} - \frac{|\nabla \psi|^2}{2} - \mu^2 c^2 \frac{|\psi|^2}{2},$$

$$(2.13) \quad \mathcal{L}_S^0(x, \psi, \nabla\psi) = i \nabla_0 \psi \cdot \psi - \frac{1}{2\mu} |\nabla \psi|^2$$

Exercise 2.5 Check the Euler-Lagrange form (2.9) for the Klein-Gordon and Schrödinger equations (2.10), (2.11).

Solution

I For the Klein-Gordon equation (2.10): by Formulas (2.4) we get

$$(2.14) \quad \pi_0(x) = \nabla_0 \psi(x), \quad \pi_k(x) = -\nabla_k \psi(x), \quad \mathcal{L}_\psi = -\mu^2 c^2 \psi.$$

Hence (2.9) is equivalent to (2.10).

II For the Schrödinger equation (2.11): by Formulas (2.4) we get

$$(2.15) \quad \pi_0(x) = -i\psi(x), \quad \pi_k(x) = -\frac{1}{\mu} \nabla_k \psi(x), \quad \mathcal{L}_\psi = i \nabla_0 \psi(x).$$

Hence (2.9) is equivalent to (2.11).

2.4 The Equations with Maxwell Field

Exercise 2.6 Check the Euler-Lagrange form (2.9) for the Klein-Gordon and Schrödinger equations (2.1), (2.2).

Solution

I For the Klein-Gordon equation (2.1): by Formulas (2.4) we get

$$(2.16) \quad \begin{cases} \pi_0(x) = -i\hbar[i\hbar\nabla_0 - \frac{e}{c}\phi(x)]\psi(x), & \pi_k(x) = -i\hbar[-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x)]\psi(x), \\ \mathcal{L}_\psi = -\frac{e}{c}\phi(x)[i\hbar\nabla_0 - \frac{e}{c}\phi(x)]\psi(x) + \frac{e}{c}\mathbf{A}_k(x)[-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x)]\psi(x) - \mu^2 c^2 \psi. \end{cases}$$

Hence (2.9) is equivalent to (2.1).

II For the Schrödinger equation (2.2): by Formulas (2.4) we get

$$(2.17) \quad \begin{cases} \pi_0(x) = -i\hbar\psi(x), & \pi_k(x) = -\frac{1}{\mu}i\hbar[-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x)]\psi(x), \\ \mathcal{L}_\psi = i\hbar\nabla_0\psi(x) - 2e\phi(x)\psi(x) + \frac{e}{\mu c}\mathbf{A}_k(x)[-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x)]\psi(x). \end{cases}$$

Hence (2.9) is equivalent to (2.2).

2.5 Action Functional

Definition 2.7 For $k = 1, 2, \dots$ and $\sigma > 0$ the space C_σ^k is the set of the functions $\psi(t, \mathbf{x}) \in C^k([0, \infty) \times \mathbb{R}^3, \mathbb{R}^N)$ with the space-decay

$$(2.18) \quad \sum_{|\alpha| \leq k} |\nabla^\alpha \psi(t, \mathbf{x})| \leq C_T(1 + |\mathbf{x}|)^{-\sigma}, \quad (t, \mathbf{x}) \in C^k([0, T] \times \mathbb{R}^3).$$

We will consider the real-valued functionals \mathcal{F} on C_σ^1 . By definition, \mathcal{F} is a map $C_\sigma^1 \rightarrow \mathbb{R}$.

Definition 2.8 For $\psi \in C_\sigma^1$ the Gateau derivative $D\mathcal{F}(\psi)$ is the **linear functional**

$$(2.19) \quad \langle D\mathcal{F}(\psi), h \rangle = \left. \frac{d}{d\varepsilon} \mathcal{F}(\psi + \varepsilon h) \right|_{\varepsilon=0}$$

for $h(\cdot) \in C_\sigma^1$ if the derivative exists.

We will assume the following bounds for the Lagrangian density,

$$(2.20) \quad \begin{array}{l} |\mathcal{L}(x, \psi, \nabla\psi)| \leq C(|\psi| + |\nabla\psi|)^2 \\ |\mathcal{L}_\psi(x, \psi, \nabla\psi)| + |\mathcal{L}_{\nabla\psi}(x, \psi, \nabla\psi)| \leq C(|\psi| + |\nabla\psi|) \end{array} \quad \left|, \quad |\psi| + |\nabla\psi| \leq \text{const}.\right.$$

Let us fix a $T > 0$.

Definition 2.9 The action for the field is the functional on C_σ^1 , $\sigma > 3/2$ defined by

$$(2.21) \quad S_T(\psi) = \int_0^T \left[\int_{\mathbb{R}^3} \mathcal{L}(x, \psi(x), \nabla\psi(x)) d\mathbf{x} \right] dx_0, \quad \psi(\cdot) \in C_\sigma^1.$$

Note that for $\sigma > 3/2$ the action is defined on the whole of C_σ^1 by (2.18) and the first inequality of (2.20). Moreover, the functional is differentiable:

Lemma 2.10 *The Gateau derivative $\langle DS_T(\psi), h \rangle$ exists for $\psi, h \in C_\sigma^1$ if $\sigma > 3/2$.*

Proof From Definition 2.8 we get by the theorem of the differentiation of integrals,

$$(2.22) \quad \begin{aligned} \langle DS_T(\psi), h \rangle &:= \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int_0^T \left[\int_{\mathbb{R}^3} \mathcal{L}(x, \psi(x) + \varepsilon h(x), \nabla \psi(x) + \varepsilon \nabla h(x)) d\mathbf{x} \right] dx_0 \\ &= \int_0^T \left[\int_{\mathbb{R}^3} \left(\mathcal{L}_\psi(x, \psi(x), \nabla \psi(x)) h(x) + \sum_{\alpha=0}^3 \mathcal{L}_{\nabla_\alpha \psi}(x, \psi(x), \nabla \psi(x)) \nabla_\alpha h(x) \right) d\mathbf{x} \right] dx_0 \end{aligned}$$

since the integrals converge uniformly by (2.18) and (2.20). ■

2.6 Hamilton Least Action Principle

Let us introduce the space of *variations*.

Definition 2.11 $C^1(T)$ is the space of functions $h(\cdot) \in C^1([0, T] \times \mathbb{R}^3, \mathbb{R}^N)$ such that

$$(2.23) \quad h(0, \mathbf{x}) = h(T, \mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^3,$$

$$(2.24) \quad h(t, \mathbf{x}) = 0, \quad |\mathbf{x}| \geq \text{const}, \quad t \in [0, T].$$

Definition 2.12 The function $\psi \in C_\sigma^1$ satisfies the Hamilton Least Action Principle (LAP) if for any $T > 0$,

$$(2.25) \quad \langle DS_T(\psi), h \rangle = 0, \quad \forall h(\cdot) \in C^1(T).$$

Theorem 2.13 For $\psi \in C_\sigma^2$ with $\sigma > 3/2$ the Hamilton LAP is equivalent to the Euler-Lagrange equations (2.7).

Proof (2.22) implies by partial integration in x_α , $\alpha = 0, \dots, 3$,

$$(2.26) \quad \begin{aligned} &\langle DS_T(\psi), h \rangle \\ &= \int_0^T \left[\int_{\mathbb{R}^3} \left(\mathcal{L}_\psi(x, \psi(x), \nabla \psi(x)) h(x) - \sum_{\alpha=0}^3 \nabla_\alpha \mathcal{L}_{\nabla_\alpha \psi}(x, \psi(x), \nabla \psi(x)) \right) h(x) d\mathbf{x} \right] dx_0. \end{aligned}$$

Therefore, (2.25) is equivalent to (2.7) by the main lemma of the calculus of variations. ■

3 Four Conservation Laws for Lagrangian Fields

By the general Noether theorem 39.9, an invariance of the Lagrangian density of a field with respect to a *symmetry group* provides a conservation law. In particular, the theorem implies four classical conservation laws: conservation of total energy, momentum, angular momentum and charge. Here, the energy conservation follows from an invariance under time translations, the momentum conservation follows from translation invariance in space, the angular momentum conservation follows from rotation invariance, and the charge conservation follows from a phase invariance.

Here we state the four conservation laws. All these conservation laws are deduced from the Noether theorem of Section 27.2. Let us fix a $\sigma > 3/2$.

3.1 Time Invariance: Energy Conservation

Definition 3.1 *The energy of the Lagrangian field $\psi(x) \in C_\sigma^1$ is defined by (cf. (36.5), (36.17))*

$$(3.1) \quad E(t) = \int_{\mathbb{R}^3} \left[\pi_0(x) \nabla_0 \psi(x) - \mathcal{L}(x, \psi(x), \nabla \psi(x)) \right] d\mathbf{x}, \quad t \in \mathbb{R}.$$

Remark 3.2 *By definition, $t = x_0$ and $\pi_0(x) \nabla_k \psi(x) = \pi_0^j(x) \nabla_k \psi^j(x)$.*

Note that by (2.18) and (2.20) the energy and momentum are well defined for the solutions $\psi(t) \in C_\sigma^1$ since $\sigma > 3/2$.

Example 3.3 *The linear wave equation*

$$(3.2) \quad \ddot{\psi}(t, \mathbf{x}) = \Delta \psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3$$

is a particular case of the Klein-Gordon equation (2.1) with $m = 0$. The Lagrangian density \mathcal{L} is given by (2.5), i.e.

$$(3.3) \quad \mathcal{L}(t, \mathbf{x}, \psi, \dot{\psi}, \psi') = \frac{|\dot{\psi}|^2}{2} - \frac{|\nabla \psi|^2}{2}.$$

Then the canonically conjugate field is $\pi_0 = \dot{\psi}$. Hence the energy is given by

$$(3.4) \quad E(t) = \int_{\mathbb{R}^3} \left[\frac{|\dot{\psi}(t, \mathbf{x})|^2}{2} + \frac{|\nabla \psi(t, \mathbf{x})|^2}{2} \right] d\mathbf{x}, \quad t \in \mathbb{R}.$$

Theorem 3.4 *Let the Lagrangian density \mathcal{L} not depend on $x_0 := t$,*

$$(3.5) \quad \mathcal{L}(x, \psi, \nabla \psi) \equiv \mathcal{L}_1(\mathbf{x}, \psi, \nabla \psi).$$

Then for any trajectory $\psi(x) \in C_\sigma^2$ of the equations (2.9) the energy is conserved, $E(t) = \text{const}$.

Proof for a particular case Let us prove the theorem for an equation of type (3.2) in dimension one,

$$(3.6) \quad \ddot{\psi}(t, x) = \psi''(t, x), \quad x \in \mathbb{R}.$$

First, $E(t) = \lim_{R \rightarrow \infty} E_R(t)$ where

$$(3.7) \quad E_R(t) = \int_{-R}^R \left[\frac{|\dot{\psi}(t, x)|^2}{2} + \frac{|\psi'(t, x)|^2}{2} \right] dx.$$

Differentiating, we get

$$(3.8) \quad \dot{E}_R(t) = \int_{-R}^R \left[\dot{\psi}(t, x) \ddot{\psi}(t, x) + \psi'(t, x) \dot{\psi}'(t, x) \right] dx.$$

Substituting here $\ddot{\psi}(t, x) = \psi''(t, x)$, we get by partial integration

$$(3.9) \quad \dot{E}_R(t) = \left[\dot{\psi}(t, x) \psi'(t, x) \right] \Big|_{-R}^R = \dot{\psi}(t, R) \psi'(t, R) - \dot{\psi}(t, -R) \psi'(t, -R).$$

Now, Eq. (2.18) implies that for every fixed t we have $\dot{E}_R(t) \rightarrow 0$ as $R \rightarrow \infty$, hence $E(t) = \text{const.}$ Indeed, for any $T > 0$,

$$(3.10) \quad E_R(T) - E_R(0) = \int_0^T \dot{E}(t) dt,$$

hence in the limit $R \rightarrow \infty$,

$$(3.11) \quad E(T) - E(0) = \int_0^T \lim_{R \rightarrow \infty} \dot{E}_R(t) dt = 0. \quad \blacksquare$$

Proof for the general case By Definition (3.1), $E(t) = \lim_{R \rightarrow \infty} E_R(t)$ where

$$(3.12) \quad E_R(t) = \int_{|\mathbf{x}| < R} \left[\pi_0(x) \nabla_0 \psi(x) - \mathcal{L}(x, \psi(x), \nabla \psi(x)) \right] \Big|_{x_0=ct} d\mathbf{x}.$$

Differentiating, we get

$$(3.13) \quad \begin{aligned} \dot{E}_R(t) &= \int_{|\mathbf{x}| < R} \left[\nabla_0 \pi_0(x) \nabla_0 \psi(x) + \pi_0(x) \nabla_0^2 \psi(x) \right. \\ &\quad \left. - \mathcal{L}_\psi(x, \psi(x), \nabla \psi(x)) \nabla_0 \psi(x) - \sum_{\alpha=0}^3 \pi_\alpha(x) \nabla_0 \nabla_\alpha \psi \right] d\mathbf{x}. \end{aligned}$$

By (2.9), we have $\nabla_0 \pi_0(x) = \mathcal{L}_\psi(x, \psi(x), \nabla \psi(x)) - \sum_1^3 \nabla_k \pi_k(x)$. Substituting into (3.13), we get by the Stokes theorem,

$$(3.14) \quad \begin{aligned} \dot{E}_R(t) &= - \int_{|\mathbf{x}| < R} \left[\sum_1^3 \nabla_k \pi_k(x) \nabla_0 \psi + \sum_{\alpha=1}^3 \pi_\alpha(x) \nabla_0 \nabla_\alpha \psi \right] d\mathbf{x} \\ &= - \int_{|\mathbf{x}| < R} \sum_1^3 \nabla_k \left[\pi_k(x) \nabla_0 \psi(x) \right] d\mathbf{x} \\ &= - \int_{|\mathbf{x}|=R} \sum_1^3 n_k(\mathbf{x}) \left[\pi_k(x) \nabla_0 \psi(x) \right] dS, \end{aligned}$$

where $n_k(\mathbf{x}) := x_k/|\mathbf{x}|$, and dS is the Lebesgue measure on the sphere $|\mathbf{x}| = R$. Now, Eq. (2.18) implies that for every fixed t we have $\dot{E}_R(t) \rightarrow 0$ as $R \rightarrow \infty$, hence $E(t) = \text{const.}$ \blacksquare

Remark 3.5 *The identity (3.14) means that the vector-function*

$$(3.15) \quad S_k(x) := \pi_k(x) \nabla_0 \psi(x) := \pi_k^j(x) \nabla_0 \psi^j(x), \quad k = 1, \dots, 3, \quad x \in \mathbb{R}^4$$

is the energy current density.

3.2 Translation Invariance: Momentum Conservation

Definition 3.6 The momentum of the Lagrangian field at time t is the vector

$$(3.16) \quad \mathbf{p}_n(t) = - \int_{\mathbb{R}^3} \left[\pi_0(x) \nabla_k \psi(x) \right] d\mathbf{x}, \quad n = 1, 2, 3, \quad t \in \mathbb{R}.$$

Example 3.7 For the linear wave equation (3.2), the momentum is given by

$$(3.17) \quad \mathbf{p}(t) = - \int_{\mathbb{R}^3} \dot{\psi}(t, \mathbf{x}) \nabla \psi(t, \mathbf{x}) d\mathbf{x}, \quad t \in \mathbb{R}.$$

Theorem 3.8 Assume that the Lagrangian density $\mathcal{L}(x, \psi, \nabla \psi)$ does not depend on the coordinate x_n . Then for any trajectory $\psi(x) \in C_\sigma^2$ of Eq. (2.9) the corresponding component of momentum is conserved, $\mathbf{p}_n(t) = \text{const}$.

Proof for a particular case Let us prove the theorem for the d'Alembert equation (3.6). By definition, $P(t) = \lim_{R \rightarrow \infty} P_R(t)$ where

$$(3.18) \quad P_R(t) = - \int_{-R}^R \dot{\psi}(t, x) \psi'(t, x) dx.$$

Differentiating, we get

$$(3.19) \quad \dot{P}_R(t) = - \int_{-R}^R \left[\ddot{\psi}(t, x) \psi'(t, x) + \dot{\psi}(t, x) \dot{\psi}'(t, x) \right] dx.$$

Substituting here $\ddot{\psi}(t, x) = \psi''(t, x)$, we get,

$$(3.20) \quad \dot{P}_R(t) = - \left[\frac{|\psi'(t, x)|^2}{2} + \frac{|\dot{\psi}(t, x)|^2}{2} \right] \Big|_{-R}^R.$$

Now, Eq. (2.18) implies that, for every fixed t , we have $\dot{P}_R(t) \rightarrow 0$ as $R \rightarrow \infty$, hence $P(t) = \text{const}$. ■

Exercise 3.9 Prove Theorem 3.8 for the general case. **Hint:** Differentiate in time and apply partial integration in \mathbf{x} .

3.3 Rotation Invariance: Angular Momentum Conservation

Let us denote by $O_n(s)$ the rotations of the space \mathbb{R}^3 around a unit vector $\mathbf{e}_n \in \mathbb{R}^3$ ($\mathbf{e}_1 := (1, 0, 0)$, etc), with an angle of s radian. Let us consider the Lagrangian densities which are invariant with respect to rotations:

$$(3.21) \quad \mathcal{L}(x_0, \mathbf{x}, \psi, \nabla_0 \psi, \nabla_{\mathbf{x}} \psi) \equiv \mathcal{L}(x_0, O_n(-s)\mathbf{x}, \psi, \nabla_0 \psi, O_n^t(s) \nabla_{\mathbf{x}} \psi).$$

Example 3.10 Let us consider the case of $\mathbf{e}_3 := (0, 0, 1)$. Then the invariance (3.21) holds for Lagrangian densities of the following structure:

$$(3.22) \quad \mathcal{L}(x, \psi(x), \nabla \psi(x)) \equiv \mathcal{L}_*(x_0, |(x_1, x_2)|, x_3, \psi(x), \nabla_0 \psi(x), |(\nabla_1 \psi(x), \nabla_2 \psi(x))|, \nabla_3 \psi(x)).$$

Definition 3.11 The angular momentum of the Lagrangian field at time t is vector

$$(3.23) \quad \mathbf{J}_n(t) = \int_{\mathbb{R}^d} \left[\pi_0(x) (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x) \right] d\mathbf{x}, \quad n = 1, 2, 3, \quad t \in \mathbb{R},$$

where $\nabla_{\mathbf{x}} := (\nabla_1, \nabla_2, \nabla_3)$: for example, $(\mathbf{x} \times \nabla_{\mathbf{x}})_3 := x_1 \nabla_2 - x_2 \nabla_1$ etc.

Note that the integral converges for solutions $\psi(x) \in C_\sigma^1$ with $\sigma > 2$, by (2.18) and (2.20).

Remark 3.12 By definition, $\pi_0(x) (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x) = \pi_0^j(x) (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi^j(x)$.

Theorem 3.13 Let the Lagrangian density satisfy (3.22) and $\sigma > 5/2$. Then, for any trajectory $\psi(x) \in C_\sigma^2$ of Eq. (2.9), the corresponding component of angular momentum is conserved, $\mathbf{J}_n(t) = \text{const}$.

Exercise 3.14 Prove Theorem 3.13.

Exercise 3.15 Calculate the angular momentum for the Klein-Gordon and Schrödinger equations (2.1), (2.2).

3.4 Phase Invariance: Charge Conservation

Consider complex-valued fields ψ with Lagrangian densities which are invariant with respect to rotations in ψ :

$$(3.24) \quad \mathcal{L}(x, e^{i\theta}\psi(x), e^{i\theta}\nabla\psi(x)) \equiv \mathcal{L}(x, \psi(x), \nabla\psi(x)), \quad \theta \in \mathbb{R}.$$

Exercise 3.16 Check that (3.24) holds for Eqns (2.1) and (2.2).

Definition 3.17 The charge of the Lagrangian field at time t is defined by

$$(3.25) \quad Q(t) = \int_{\mathbb{R}^d} [\pi_0(x) \cdot i\psi(x)] d\mathbf{x}, \quad t \in \mathbb{R},$$

where $\pi_0(x)$ and $i\psi(x)$ are identified with real vectors in \mathbb{R}^2 , and \cdot is the real scalar product in \mathbb{R}^2 (cf. (2.4)).

Note that the integral converges for the solutions $\psi(t) \in C_\sigma^1$ by (2.18) and (2.20) since $\sigma > 3/2$.

Theorem 3.18 Let the Lagrangian density satisfy (3.24) and $\sigma > 3/2$. Then for any trajectory $\psi(x) \in C_\sigma^2$ of Eq. (2.9) the charge is conserved, $Q(t) = \text{const}$.

Proof for a particular case Let us prove the theorem for the free linear 1D Klein-Gordon equation (2.1),

$$(3.26) \quad \ddot{\psi}(t, x) = \psi''(t, x) - \mu^2\psi(t, x), \quad \mathbf{x} \in \mathbb{R}.$$

Then $\pi = \dot{\psi}$ and $Q(t) = \lim_{R \rightarrow \infty} Q_R(t)$, where

$$(3.27) \quad Q_R(t) = \int_{-R}^R \dot{\psi}(t, x) \cdot i\psi(t, x) dx.$$

Differentiating, we get

$$(3.28) \quad \dot{Q}_R(t) = \int_{-R}^R \ddot{\psi}(t, x) \cdot i\psi(t, x) dx + \int_{-R}^R \dot{\psi}(t, x) \cdot i\dot{\psi}(t, x) dx.$$

The second integral on the RHS is zero since $z_1 \cdot iz_2$ is an antisymmetric bilinear form in \mathbb{C} by (2.4). Hence, substituting $\ddot{\psi}(t, x) = \psi''(t, x) - \mu^2\psi(t, x)$ from (3.26), we get by partial integration,

$$(3.29) \quad \begin{aligned} \dot{Q}_R(t) &= \psi'(t, x) \cdot i\psi(t, x) \Big|_{-R}^R - \int_{-R}^R \psi'(t, x) \cdot i\psi'(t, x) dx - \mu^2 \int_{-R}^R \psi(t, x) \cdot i\psi(t, x) dx \\ &= \psi'(t, x) \cdot i\psi(t, x) \Big|_{-R}^R = j(t, -R) - j(t, R), \quad j(t, x) := -\psi'(t, x) \cdot i\psi(t, x). \end{aligned}$$

since both integrals are zero by antisymmetry. Therefore, Eq. (2.18) implies that for every fixed t we have $\dot{Q}_R(t) \rightarrow 0$ as $R \rightarrow \infty$, hence $Q(t) = \text{const}$. ■

Exercise 3.19 Prove Theorem 3.18 for the nonlinear 1D Schrödinger equation (2.2),

$$(3.30) \quad i\dot{\psi}(t, \mathbf{x}) = \frac{1}{2\mu}\psi''(t, \mathbf{x}) + F(\psi), \quad \mathbf{x} \in \mathbb{R},$$

where $F(\psi) = -\nabla U(|\psi|)$, $\psi \in \mathbb{C} \equiv \mathbb{R}^2$.

Exercise 3.20 Prove Theorem 3.18 for the general case and check the **continuity equation**,

$$(3.31) \quad \dot{\rho}(t, x) + \operatorname{div} \mathbf{j}(t, x) = 0,$$

where $\rho(t, x) := \pi_0(x) \cdot i\psi(x)$ is the **charge density** and $\mathbf{j}_k(t, x) := \pi_k(t, \mathbf{x}) \cdot i\psi(t, \mathbf{x})$ is the **current density**.

Exercise 3.21 Calculate the charge- and current densities for the Klein-Gordon and Schrödinger equations (2.1), (2.2).

4 Lagrangian Theory for the Maxwell Field

The evolution of the electromagnetic field is governed by the Maxwell equations, which describe, among other features, the propagation of light. The Maxwell equations already obey the symmetries of special relativity. The formulation of the Maxwell equations in terms of *potentials* rather than fields is especially useful, because these potentials show up in the matter equations (e.g., in the Schrödinger and Klein-Gordon equations).

The Maxwell equations, as well, can be derived from a least action principle, and, therefore, a Lagrangian formulation of the Maxwell theory may be given. A classical, point-like charged particle in an electromagnetic field is described by the non-relativistic or relativistic Lorentz equations, respectively. Again, a Lagrangian and Hamiltonian formulation for the Lorentz equations can be found.

4.1 Maxwell Equations and Potentials. Lagrangian Density

In 1862, Maxwell completed the Coulomb, Faraday and Biot-Savart-Laplace equations by the *displacement current* and wrote the complete system of *classical electrodynamics*. In *Gaussian units* it reads

$$(4.1) \quad \left\{ \begin{array}{l} \operatorname{div} \mathbf{E}(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}), \quad \operatorname{rot} \mathbf{E}(t, \mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0, \quad \operatorname{rot} \mathbf{B}(t, \mathbf{x}) = \frac{4\pi}{c}\mathbf{j}(t, \mathbf{x}) + \frac{1}{c}\dot{\mathbf{E}}(t, \mathbf{x}), \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

where $\rho(t, \mathbf{x})$ and $\mathbf{j}(t, \mathbf{x})$ stand for the charge and current density, respectively.

Remark 4.1 *The Maxwell equations imply the charge continuity equation,*

$$(4.2) \quad \dot{\rho}(t, \mathbf{x}) + \operatorname{div} \mathbf{j}(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Potentials and gauge invariance

Let us introduce scalar and vector potentials to rewrite (4.1) in a relativistic covariant four-dimensional form. Namely, $\operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0$ implies that $\mathbf{B}(t, \mathbf{x}) = \operatorname{rot} \mathbf{A}(t, \mathbf{x})$. Then $\operatorname{rot} \mathbf{E}(t, \mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}(t, \mathbf{x})$ implies $\operatorname{rot} [\mathbf{E}(t, \mathbf{x}) + \frac{1}{c}\dot{\mathbf{A}}(t, \mathbf{x})] = 0$ hence $\mathbf{E}(t, \mathbf{x}) + \frac{1}{c}\dot{\mathbf{A}}(t, \mathbf{x}) = -\nabla_{\mathbf{x}}\phi(t, \mathbf{x})$. Finally,

$$(4.3) \quad \mathbf{B}(t, \mathbf{x}) = \operatorname{rot} \mathbf{A}(t, \mathbf{x}), \quad \mathbf{E}(t, \mathbf{x}) = -\nabla_{\mathbf{x}}\phi(t, \mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

The justification of all these relations follows from the Fourier transform. The choice of the potentials is not unique since the *gauge transformation*

$$(4.4) \quad \phi(t, \mathbf{x}) \mapsto \phi(t, \mathbf{x}) + \frac{1}{c}\dot{\chi}(t, \mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) \mapsto \mathbf{A}(t, \mathbf{x}) - \nabla_{\mathbf{x}}\chi(t, \mathbf{x})$$

does not change the fields $\mathbf{E}(t, \mathbf{x})$, $\mathbf{B}(t, \mathbf{x})$ for any function $\chi(t, \mathbf{x}) \in C^1(\mathbb{R}^4)$. Therefore, it is possible to satisfy an additional *gauge* condition. Let us choose for example the *Lorentz gauge*

$$(4.5) \quad \frac{1}{c}\dot{\phi}(t, \mathbf{x}) + \operatorname{div} \mathbf{A}(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Exercise 4.2 *Prove the existence of the potentials satisfying (4.5). Hint: Derive an equation for the function $\chi(t, \mathbf{x})$.*

Let us express the Maxwell equations (4.1) in terms of the potentials. Substituting Eq. (4.3) into the first Maxwell equation leads to $4\pi\rho(t, \mathbf{x}) = \operatorname{div} \mathbf{E}(t, \mathbf{x}) = -\Delta\phi(t, \mathbf{x}) - \frac{1}{c} \operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x})$. Eliminating $\operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x})$ by the differentiation of (4.5) in time, $\frac{1}{c} \ddot{\phi}(t, \mathbf{x}) + \operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x}) = 0$, we get

$$(4.6) \quad \square\phi(t, \mathbf{x}) := \left[\frac{1}{c^2} \partial_t^2 - \Delta\right]\phi(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Similarly, substituting Eq. (4.3) into the last Maxwell equation, we get

$$(4.7) \quad \operatorname{rot} \operatorname{rot} \mathbf{A}(t, \mathbf{x}) = \frac{4\pi}{c} \mathbf{j}(t, \mathbf{x}) + \frac{1}{c} \dot{\mathbf{E}}(t, \mathbf{x}) = \frac{4\pi}{c} \mathbf{j}(t, \mathbf{x}) - \frac{1}{c} \nabla_{\mathbf{x}} \dot{\phi}(t, \mathbf{x}) - \frac{1}{c^2} \ddot{\mathbf{A}}(t, \mathbf{x}).$$

Exercise 4.3 *Prove the identity*

$$(4.8) \quad \operatorname{rot} \operatorname{rot} = -\Delta + \nabla_{\mathbf{x}} \operatorname{div} .$$

Substituting Eq. (4.8) into (4.7) and eliminating $\frac{1}{c} \nabla_{\mathbf{x}} \dot{\phi}(t, \mathbf{x})$ by application of $\nabla_{\mathbf{x}}$ to (4.5), we get

$$(4.9) \quad \square\mathbf{A}(t, \mathbf{x}) = \frac{4\pi}{c} \mathbf{j}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Remark 4.4 *The arguments above show that the Maxwell equations (4.1) are equivalent to the system of two wave equations (4.6), (4.9) for the potentials with the Lorentz gauge condition (4.5).*

4D vector potential

Let us introduce the four-dimensional notations

$$(4.10) \quad \begin{cases} x_0 = ct, & x_\mu = (x_0, \dots, x_3), & \partial_\mu = \nabla_\mu = (\partial_0, \partial_1, \partial_2, \partial_3), \\ x^\mu := g^{\mu\nu} x_\nu = (x_0, -x_1, -x_2, -x_3), & \partial^\mu := g^{\mu\nu} \partial_\nu = (\partial_0, -\partial_1, -\partial_2, -\partial_3), \end{cases}$$

where $g_{\mu\nu} = g^{\mu\nu} = \operatorname{diag}(1, -1, -1, -1)$. Let us also introduce the four-dimensional fields and currents

$$(4.11) \quad \left\{ \begin{array}{l} \mathcal{A}^\mu(x) = (\phi(t, \mathbf{x}), \mathbf{A}(x)), \quad \mathcal{A}_\mu(x) := g_{\mu\nu} \mathcal{A}^\nu(x) = (\phi(x), -\mathbf{A}(x)), \\ \mathcal{J}^\mu(x) = (\rho(x), \frac{1}{c} \mathbf{j}(x)), \quad \mathcal{J}_\mu(x) := g_{\mu\nu} \mathcal{J}^\nu(x) = (\rho(x), -\frac{1}{c} \mathbf{j}(x)). \end{array} \right. \quad x \in \mathbb{R}^4$$

Then the Maxwell equations (4.6), (4.9) become

$$(4.12) \quad \square\mathcal{A}^\mu(x) = 4\pi\mathcal{J}^\mu(x), \quad x \in \mathbb{R}^4.$$

Similarly, the charge continuity equation (4.2), gauge transformation (4.4) and Lorentz gauge (4.5) become

$$(4.13) \quad \partial_\mu \mathcal{J}^\mu(x) = 0, \quad \mathcal{A}_\mu(x) \mapsto \mathcal{A}_\mu(x) + \partial^\mu \chi(x), \quad \partial_\mu \mathcal{A}^\mu(x) = 0, \quad x \in \mathbb{R}^4.$$

Tensor field

Definition 4.5 *The Maxwell tensor is defined by*

$$(4.14) \quad \mathcal{F}^{\mu\nu}(x) = \partial^\mu \mathcal{A}^\nu(x) - \partial^\nu \mathcal{A}^\mu(x), \quad x \in \mathbb{R}^4.$$

Exercise 4.6 Check the formula

$$(4.15) \quad (\mathcal{F}^{\mu\nu}(x))_{\mu,\nu=0}^3 = \begin{pmatrix} 0 & -\mathbf{E}_1(x) & -\mathbf{E}_2(x) & -\mathbf{E}_3(x) \\ \mathbf{E}_1(x) & 0 & -\mathbf{B}_3(x) & \mathbf{B}_2(x) \\ \mathbf{E}_2(x) & \mathbf{B}_3(x) & 0 & -\mathbf{B}_1(x) \\ \mathbf{E}_3(x) & -\mathbf{B}_2(x) & \mathbf{B}_1(x) & 0 \end{pmatrix}$$

Hint: Use the formulas (4.3) and (4.10), (4.11).

Proposition 4.7 The Maxwell equations (4.12) are equivalent to

$$(4.16) \quad \partial_\mu \mathcal{F}^{\mu\nu}(x) = 4\pi \mathcal{J}^\nu(x), \quad x \in \mathbb{R}^4.$$

Proof The Maxwell tensor does not depend on the choice of gauge since $\partial^\mu \partial^\nu \chi(x) - \partial^\nu \partial^\mu \chi(x) = 0$. Therefore, we can assume the Lorentz gauge (4.5) without loss of generality. Then (4.12) implies

$$(4.17) \quad \partial_\mu \mathcal{F}^{\mu\nu}(x) = \partial_\mu (\partial^\mu \mathcal{A}^\nu(x) - \partial^\nu \mathcal{A}^\mu(x)) = \partial_\mu \partial^\mu \mathcal{A}^\nu(x) = \square \mathcal{A}^\nu(x) = 4\pi \mathcal{J}^\nu(x), \quad x \in \mathbb{R}^4. \quad \blacksquare$$

Lagrangian density

Definition 4.8 The Lagrangian density for the Maxwell equations (4.16) with **given** external charge-current densities $\mathcal{J}^\nu(x)$ is defined by

$$(4.18) \quad \mathcal{L}(x, \mathcal{A}_\mu, \nabla \mathcal{A}_\mu) = -\frac{1}{16\pi} \mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu} - \mathcal{J}^\nu(x) \mathcal{A}_\nu, \quad (x, \mathcal{A}_\mu, \nabla \mathcal{A}_\mu) \in \mathbb{R}^4 \times \mathbb{R}^4 \times \mathbb{R}^{16}.$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$.

Proposition 4.9 The Maxwell equations (4.16) with given charge-current densities $\mathcal{J}^\nu(x)$ are equivalent to the Euler-Lagrange equations (2.9) with the Lagrangian density (4.18) for the fields $\mathcal{A}_\mu(x)$.

Proof $\mathcal{F}_{\mu\nu} \mathcal{F}^{\mu\nu}$ is a quadratic form in $\nabla \mathcal{A}$. Therefore, the canonically conjugate fields π_α with components $\pi_{\alpha\beta}$ are given by

$$(4.19) \quad \pi_{\alpha\beta} := \nabla_{\partial_\alpha \mathcal{A}_\beta} \mathcal{L} = -\frac{1}{8\pi} \mathcal{F}^{\mu\nu} \nabla_{\partial_\alpha \mathcal{A}_\beta} \mathcal{F}_{\mu\nu} = -\frac{1}{8\pi} (\mathcal{F}^{\alpha\beta} - \mathcal{F}^{\beta\alpha}) = -\frac{1}{4\pi} \mathcal{F}^{\alpha\beta}.$$

Therefore,

$$(4.20) \quad \nabla_\alpha \pi_{\alpha\beta}(x) = -\frac{1}{4\pi} \nabla_\alpha \mathcal{F}^{\alpha\beta}(x), \quad x \in \mathbb{R}^4.$$

On the other hand, $\partial_{\mathcal{A}_\beta} \mathcal{L} = -\mathcal{J}_\beta(x)$, hence (4.16) is equivalent to (2.9). ■

4.2 Lagrangian for Charged Particle in Maxwell Field

In a **continuous** Maxwell field the motion of a charged particle with small velocity $|\dot{\mathbf{x}}(t)| \ll c$ is governed by the Lorentz equation (L) from the Introduction,

$$(4.21) \quad \mu \ddot{\mathbf{x}}(t) = e[\mathbf{E}(t, \mathbf{x}(t)) + \frac{1}{c} \dot{\mathbf{x}}(t) \times \mathbf{B}(t, \mathbf{x}(t))], \quad t \in \mathbb{R},$$

where μ is the mass of the particle and e its charge. For large velocities $|\dot{\mathbf{x}}(t)| \sim c$ the equation must be replaced by

$$(4.22) \quad \dot{\mathbf{p}}(t) = e[\mathbf{E}(t, \mathbf{x}(t)) + \frac{1}{c} \dot{\mathbf{x}}(t) \times \mathbf{B}(t, \mathbf{x}(t))], \quad t \in \mathbb{R},$$

where $\mathbf{p} := \mu \dot{\mathbf{x}} / \sqrt{1 - (\dot{\mathbf{x}}/c)^2}$ is the relativistic momentum of the particle (see (38.6)). Let us assume that the fields \mathbf{E}, \mathbf{B} are C^1 vector functions in \mathbb{R}^4 . Then the dynamical equations (4.21), (4.22) define the corresponding dynamics uniquely. Let us define the charge-current densities $\mathcal{J}^\nu(t, \mathbf{x})$ corresponding to the trajectory $\mathbf{x}(\cdot)$: by (4.11),

$$(4.23) \quad \mathcal{J}^0(t, \mathbf{x}) = e\delta(x - \mathbf{x}(t)), \quad \mathcal{J}^k(t, \mathbf{x}) = \frac{1}{c} \dot{\mathbf{x}}_k(t) e\delta(x - \mathbf{x}(t)), \quad k = 1, 2, 3.$$

Let us show that the dynamical equations (4.21), (4.22) *automatically follow* from the Hamilton LAP applied to the Lagrangian density (4.18) with fixed fields $E(t, \mathbf{x}), B(t, \mathbf{x})$. We consider (4.21) for concreteness. Let $\mathcal{A}(t, \mathbf{x})$ be the 4-potential corresponding to the fields $\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x})$. The Lagrangian density (4.18) consists of two parts: the field part \mathcal{L}_f and the *field-matter interaction* part \mathcal{L}_{fm} . Substituting (4.23) into \mathcal{L}_{fm} , we get the field-matter action in the form

$$(4.24) \quad \begin{aligned} S_{fm}^T &= \int_0^T \int_{\mathbb{R}^3} \mathcal{J}^\nu(t, \mathbf{x}) \mathcal{A}_\nu(t, \mathbf{x}) d\mathbf{x} dt \\ &= e \int_0^T \left[\phi(t, \mathbf{x}(t)) - \frac{1}{c} \dot{\mathbf{x}}(t) \cdot \mathbf{A}(t, \mathbf{x}(t)) \right] dt. \end{aligned}$$

The interaction term corresponds to the following Lagrangian function for the nonrelativistic particle,

$$(4.25) \quad L(\mathbf{x}, \mathbf{v}, t) = \frac{\mu \mathbf{v}^2}{2} - e\phi(t, \mathbf{x}) + \frac{e}{c} \mathbf{v} \cdot \mathbf{A}(t, \mathbf{x}).$$

Theorem 4.10 *Let the potential $\mathcal{A}(t, \mathbf{x})$ be fixed. Then the Lorentz equation (4.21) for the trajectory $\mathbf{x}(\cdot)$ is equivalent to the Euler-Lagrange equations corresponding to the Lagrangian (4.25)*

Proof First let us evaluate the momentum. By definition, $\mathbf{p} := L_{\mathbf{v}} = \mu \mathbf{v} + \frac{e}{c} \mathbf{A}(t, \mathbf{x})$, hence

$$(4.26) \quad \mathbf{p}(t) := L_{\mathbf{v}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) = \mu \dot{\mathbf{x}}(t) + \frac{e}{c} \mathbf{A}(t, \mathbf{x}(t)).$$

Now (36.15) becomes,

$$(4.27) \quad \dot{\mathbf{p}}_k(t) = L_{\mathbf{x}_k}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) = -e\nabla_k \phi(t, \mathbf{x}) + \frac{e}{c} \dot{\mathbf{x}} \cdot \nabla_k \mathbf{A}(t, \mathbf{x}), \quad k = 1, 2, 3.$$

Let us calculate the derivative on the LHS:

$$(4.28) \quad \dot{\mathbf{p}}_k(t) = \frac{d}{dt} (\mu \dot{\mathbf{x}}_k(t) + \frac{e}{c} \mathbf{A}_k(t, \mathbf{x}(t))) = \mu \ddot{\mathbf{x}}_k + \frac{e}{c} [\dot{\mathbf{A}}_k(t, \mathbf{x}) + \nabla_j \mathbf{A}_k(t, \mathbf{x}) \dot{\mathbf{x}}_j].$$

Substituting this expression into the LHS of (4.27), we get

$$(4.29) \quad \mu \ddot{\mathbf{x}}_k + \frac{e}{c} [\dot{\mathbf{A}}_k(t, \mathbf{x}) + \nabla_j \mathbf{A}_k(t, \mathbf{x}) \dot{\mathbf{x}}_j] = -e\nabla_k \phi(t, \mathbf{x}) + \frac{e}{c} \dot{\mathbf{x}}_j \nabla_k \mathbf{A}_j(t, \mathbf{x}).$$

Let us rewrite this as follows:

$$(4.30) \quad \mu \ddot{\mathbf{x}}_k = e[-\nabla_k \phi(t, \mathbf{x}) - \frac{1}{c} \dot{\mathbf{A}}_k(t, \mathbf{x})] + \frac{e}{c} \dot{\mathbf{x}}_j [\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k].$$

The first square bracket on the RHS is $\mathbf{E}(t, \mathbf{x})$ by (4.3). Hence, it remains to check that $\dot{\mathbf{x}}_j [\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k] = \dot{\mathbf{x}} \times \text{rot } \mathbf{A}(t, \mathbf{x})$. Let us note that $\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k = (\text{rot } \mathbf{A})_l \varepsilon_{kjl}$ where ε_{kjl} is the antisymmetric tensor. Therefore, $\dot{\mathbf{x}}_j [\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k] = \dot{\mathbf{x}}_j (\text{rot } \mathbf{A})_l \varepsilon_{kjl} = [\dot{\mathbf{x}} \times \text{rot } \mathbf{A}(t, \mathbf{x})]_k$ by definition of the vector product. ■

Remark 4.11 *The derivation of the expression for the Lorentz force from the Maxwell equations is not very surprising since the expression also follows from the Coulomb and Biot-Savart-Laplace equations.*

4.3 Hamiltonian for Charged Particle in Maxwell Field

Nonrelativistic particle

We evaluate the Hamilton function as the Legendre transform of the nonrelativistic Lagrangian (4.25): first,

$$(4.31) \quad H := \mathbf{p}\mathbf{v} - L = \mathbf{p}\mathbf{v} - \frac{\mu\mathbf{v}^2}{2} + e\phi(t, \mathbf{x}) - \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(t, \mathbf{x}) = e\phi(t, \mathbf{x}) + \mathbf{v}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right) - \frac{\mu\mathbf{v}^2}{2}.$$

Next, we eliminate \mathbf{v} by the relation $\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x}) = \mu\mathbf{v}$. Then we get finally,

$$(4.32) \quad H = e\phi(t, \mathbf{x}) + \frac{\mu\mathbf{v}^2}{2} = e\phi(t, \mathbf{x}) + \frac{1}{2\mu}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right)^2.$$

Relativistic particle

Let us consider the relativistic Lagrangian (cf. (38.6))

$$(4.33) \quad L(\mathbf{x}, \mathbf{v}, t) = -\mu c^2 \sqrt{1 - \beta^2} - e\phi(t, \mathbf{x}) + \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(t, \mathbf{x}),$$

where $\beta := |\mathbf{v}|/c$. Let us note that the first term on the RHS is asymptotically $-\mu c^2 + \mu\mathbf{v}^2/2$ for $\beta \ll 1$. First we evaluate the momentum: by definition, $\mathbf{p} := L_{\mathbf{v}} = \mu\mathbf{v}/\sqrt{1 - \beta^2} + \frac{e}{c}\mathbf{A}(t, \mathbf{x})$, hence

$$(4.34) \quad \begin{aligned} H : &= \mathbf{p}\mathbf{v} - L = \mathbf{p}\mathbf{v} + \mu c^2 \sqrt{1 - \beta^2} + e\phi(t, \mathbf{x}) - \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(t, \mathbf{x}) \\ &= e\phi(t, \mathbf{x}) + \mathbf{v}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right) + \mu c^2 \sqrt{1 - \beta^2}. \end{aligned}$$

Next, we eliminate \mathbf{p} by the relation $\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x}) = \mu\mathbf{v}/\sqrt{1 - \beta^2}$. Then we get,

$$(4.35) \quad \begin{aligned} H &= e\phi(t, \mathbf{x}) + \frac{\mu\mathbf{v}^2}{\sqrt{1 - \beta^2}} + \mu c^2 \sqrt{1 - \beta^2} = e\phi(t, \mathbf{x}) + \frac{\mu c^2}{\sqrt{1 - \beta^2}} \\ &= e\phi(t, \mathbf{x}) + \mu c^2 \sqrt{1 + \left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right)^2 / (\mu c)^2}. \end{aligned}$$

We can rewrite this relation in the following standard form

$$(4.36) \quad \left(H/c - \frac{e}{c}\phi(t, \mathbf{x})\right)^2 = \mu^2 c^2 + \left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})\right)^2.$$

Remark 4.12 *This expression coincides with (1.10) for $\phi = 0$, $A = 0$.*

Part II

Schrödinger Equation

5 Geometric Optics and Schrödinger Equation

In this lecture we show the wave propagation along straight lines for the simple cases of free Klein-Gordon and Schrödinger equations. For the Schrödinger equation coupled to the Maxwell field we analyze the propagation along rays. The proof is based on constructing a formal Debye expansion.

Wave equations of type (4.6) describe the wave propagation in electrodynamics, acoustics and many other fields. They describe well the diffraction and the interference of waves. On the other hand, the wave processes also demonstrate the straight-line propagation of waves, thereby justifying geometric optics. The mathematical description of this feature by the wave equation has been discovered by Hamilton around 1830 and developed by Liouville in 1837, Debye in 1911, Rayleigh in 1912, Jeffreys in 1923, and Schrödinger, Wentzel, Kramers and Brillouin in 1926.

5.1 Straight Line Propagation for the Free Equations

Let us analyze the straight line propagation in the concrete example of the **free** Klein-Gordon equation

$$(5.1) \quad \left\{ \begin{array}{l} \ddot{\psi}(t, \mathbf{x}) = \Delta\psi(t, \mathbf{x}) - \mu^2\psi(t, \mathbf{x}) \\ \psi(0, \mathbf{x}) = \psi_0(\mathbf{x}), \quad \dot{\psi}(0, \mathbf{x}) = \pi_0(\mathbf{x}) \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3.$$

Let us choose the initial data $\psi_0(\mathbf{x}), \pi_0(\mathbf{x})$ from the *Schwartz space* $\mathcal{S}(\mathbb{R}^3)$ of test functions.

Definition 5.1 $\mathcal{S}(\mathbb{R}^3)$ is the space of functions $\psi(\mathbf{x}) \in C^\infty(\mathbb{R}^3)$ such that

$$(5.2) \quad \sup_{\mathbf{x} \in \mathbb{R}^3} (1 + |\mathbf{x}|)^N |\nabla_{\mathbf{x}}^\alpha \psi(\mathbf{x})| < \infty$$

for any $N = 1, 2, \dots$ and multiindices $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.

Definition 5.2 For $\psi \in \mathcal{S}(\mathbb{R}^3)$ the Fourier transform is defined by

$$(5.3) \quad F\psi(\mathbf{k}) := \hat{\psi}(\mathbf{k}) := (2\pi)^{-3} \int_{\mathbb{R}^3} e^{i\mathbf{k}\mathbf{x}} \psi(\mathbf{x}) d\mathbf{x}, \quad \mathbf{k} \in \mathbb{R}^3.$$

Proposition 5.3 Let $\psi_0, \pi_0 \in \mathcal{S}(\mathbb{R}^3)$. Then the Cauchy problem (5.1) admits a unique solution $\psi(t, \mathbf{x})$ satisfying the bounds

$$(5.4) \quad \sup_{\mathbf{x} \in \mathbb{R}^3} (1 + |\mathbf{x}|)^N |\partial_t^{\alpha_0} \nabla_{\mathbf{x}}^\alpha \psi(t, \mathbf{x})| < C(\alpha_0, \alpha, N)(1 + |t|)^N, \quad t \in \mathbb{R}$$

for any $N = 1, 2, \dots$, $\alpha_0 = 0, 1, 2, \dots$ and multiindices $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.

Proof Let us calculate the solution to (5.1) with $\psi_0, \pi_0 \in \mathcal{S}(\mathbb{R}^3)$ by using the Fourier transform. Let us apply the transform to Equations (5.1) using the well-known formulas

$$(5.5) \quad (F\partial_1\psi)(t, \mathbf{k}) = -ik_1 \hat{\psi}(t, \mathbf{k}), \quad (F\Delta\psi)(t, \mathbf{k}) = -k^2 \hat{\psi}(t, \mathbf{k}), \quad (t, \mathbf{k}) \in \mathbb{R} \times \mathbb{R}^3.$$

The bounds (5.4) imply also that $(F\dot{\psi})(t, \mathbf{k}) = \dot{\hat{\psi}}(t, \mathbf{k})$ and $(F\ddot{\psi})(t, \mathbf{k}) = \ddot{\hat{\psi}}(t, \mathbf{k})$ hence (5.1) becomes

$$(5.6) \quad \left\{ \begin{array}{l} \ddot{\hat{\psi}}(t, \mathbf{k}) = -k^2 \hat{\psi}(t, \mathbf{k}) - \mu^2 \hat{\psi}(t, \mathbf{k}) \\ \hat{\psi}(0, \mathbf{k}) = \hat{\psi}_0(\mathbf{k}), \quad \dot{\hat{\psi}}(0, \mathbf{k}) = \hat{\pi}_0(\mathbf{k}) \end{array} \right. \quad (t, \mathbf{k}) \in \mathbb{R} \times \mathbb{R}^3.$$

This is the Cauchy problem for an ordinary differential equation which depends on the parameter $\mathbf{k} \in \mathbb{R}^3$. The solution is well-known,

$$(5.7) \quad \hat{\psi}(t, \mathbf{k}) = \hat{\psi}_0(\mathbf{k}) \cos \omega t + \hat{\pi}_0(\mathbf{k}) \frac{\sin \omega t}{\omega}, \quad \omega = \omega(\mathbf{k}) := \sqrt{\mathbf{k}^2 + \mu^2}.$$

Therefore, the solution $\psi(t, \mathbf{x})$ is given by the inverse Fourier transform,

$$(5.8) \quad \begin{aligned} \psi(t, \mathbf{x}) &= \int_{\mathbb{R}^3} e^{-i\mathbf{x}\mathbf{k}} \left[\hat{\psi}_0(\mathbf{k}) \cos \omega t + \hat{\pi}_0(\mathbf{k}) \frac{\sin \omega t}{\omega} \right] d\mathbf{x} \\ &= \frac{1}{2} \int_{\mathbb{R}^3} e^{-i\mathbf{x}\mathbf{k}} \left[e^{i\omega t} \left(\hat{\psi}_0(\mathbf{k}) + \frac{\hat{\pi}_0(\mathbf{k})}{i\omega} \right) + e^{-i\omega t} \left(\hat{\psi}_0(\mathbf{k}) - \frac{\hat{\pi}_0(\mathbf{k})}{i\omega} \right) \right] d\mathbf{k} \\ &= \psi_+(t, \mathbf{x}) + \psi_-(t, \mathbf{x}). \end{aligned}$$

The representation obviously implies the bounds (5.4) since $\hat{\psi}_0, \hat{\pi}_0 \in \mathcal{S}(\mathbb{R}^3)$. ■

Now let us choose the initial functions ψ_0, π_0 with a localized spectrum. Namely,

$$(5.9) \quad \text{supp } \hat{\psi}_0, \text{supp } \hat{\pi}_0 \subset B_r(\mathbf{k}_*),$$

where $B_r(\mathbf{k}_*)$ is an *open ball* with center $\mathbf{k}_* \in \mathbb{R}^3 \setminus 0$ and a small radius $r \ll |\mathbf{k}_*|$. Then the same is true for the spectra of the functions $\psi_{\pm}(t, \mathbf{x})$ by (5.8):

$$(5.10) \quad \text{supp } \hat{\psi}_{\pm}(t, \cdot) \subset B_r(\mathbf{k}_*), \quad t \in \mathbb{R}.$$

The solutions of type $\psi_{\pm}(t, \mathbf{x})$ are called *wave packets*.

Theorem 5.4 *Let $\psi_0, \pi_0 \in \mathcal{S}(\mathbb{R}^3)$, and (5.9) holds. Then the corresponding wave packets $\psi_{\pm}(t, \mathbf{x})$ are localized solutions moving with the **group velocities** $\mathbf{v}_{\pm} = \pm \nabla \omega(\mathbf{k}_*)$ in the following sense:*

i) For a constant $a > 0$ and any $N > 0$,

$$(5.11) \quad |\psi_{\pm}(t, \mathbf{x})| \leq C_N (|t| + |\mathbf{x}|)^{-N}, \quad |\mathbf{x} - \mathbf{v}_{\pm} t| > a|t|.$$

ii) For any constant $A > 0$,

$$(5.12) \quad |\psi_{\pm}(t, \mathbf{x})| \leq C(1 + |t|)^{-3/2}, \quad |\mathbf{x} - \mathbf{v}_{\pm} t| \leq A|t|.$$

Proof Let us prove the theorem for ψ_+ since for ψ_- just the same arguments hold. Let us consider the function $\psi_+(t, \mathbf{x})$ along a ray $\mathbf{x} = \mathbf{v}t$ with an arbitrary $\mathbf{v} \in \mathbb{R}^3$: by (5.8),

$$(5.13) \quad \psi_+(t, \mathbf{v}t) = \int e^{-i\phi_+(\mathbf{k})t} \Psi_+(\mathbf{k}) d\mathbf{k}.$$

Here the *phase function* is given by $\phi_+(\mathbf{k}) := \mathbf{v}\mathbf{k} - \omega(\mathbf{k})$ and the amplitude $\Psi_+(\mathbf{k}) := \hat{\psi}_0(\mathbf{k})/2 + \hat{\pi}_0(\mathbf{k})/(2i\omega(\mathbf{k}))$. Let us apply the *method of stationary phase* [32] to the integral (5.13). Then we get that the asymptotics for $t \rightarrow \infty$ depends on the existence of the critical points $\mathbf{k} \in \text{supp } \Psi_+$ of the phase function $\phi_+(\mathbf{k})$,

$$(5.14) \quad \nabla \phi_+(\mathbf{k}) = \mathbf{v} - \nabla \omega(\mathbf{k}) = 0, \quad \mathbf{k} \in \text{supp } \Psi_+.$$

In other words, $\mathbf{v} = \nabla \omega(\mathbf{k})$ with a $\mathbf{k} \in \text{supp } \Psi_+$. Now let us take into account that $\text{supp } \Psi_+ \subset B_r(\mathbf{k}_*)$ by (5.9). Then the system (5.14) admits a solution iff $\mathbf{v} \in \mathbf{V}_r(\mathbf{k}_*) := \{\nabla \omega(\mathbf{k}) : \mathbf{k} \in B_r(\mathbf{k}_*)\}$. By continuity, for some $a > 0$ we have that

$$(5.15) \quad \overline{\mathbf{V}_r(\mathbf{k}_*)} \subset B_a(\mathbf{v}_+).$$

Now let us analyze two distinct situations separately.

i) First let us consider $\mathbf{v} \notin \mathbf{V}_a(\mathbf{v}_+)$. Then the asymptotics (5.11) follows by partial integration in (5.13) with the help of an obvious identity

$$(5.16) \quad e^{-i\phi_+(\mathbf{k})t} = \frac{D}{-it} e^{-i\phi_+(\mathbf{k})t}, \quad \mathbf{k} \in \text{supp } \Psi_+,$$

where D is the following differential operator

$$(5.17) \quad D = \frac{\nabla \bar{\phi}(\mathbf{k}) \cdot \nabla}{|\nabla \phi(\mathbf{k})|^2}.$$

It is important that $\nabla \phi(\mathbf{k}) = \mathbf{v} - \nabla \omega(\mathbf{k}) \neq 0$ for $\mathbf{k} \in \text{supp } \Psi_+$ since $\mathbf{v} \notin \mathbf{V}_r(\mathbf{k}_*)$. Moreover, (5.15) implies that for an $\varepsilon > 0$,

$$(5.18) \quad |\nabla \phi(\mathbf{k})| = |\mathbf{v} - \nabla \omega(\mathbf{k})| \geq \varepsilon(1 + |\mathbf{v}|), \quad \mathbf{v} \notin \mathbf{V}_a(\mathbf{v}_+), \quad \mathbf{k} \in \text{supp } \Psi_+.$$

Applying the identity (5.16) N times in (5.13), we get by partial integration,

$$(5.19) \quad \psi_+(t, \mathbf{v}t) = (-it)^{-N} \int D^N e^{-i\phi_+(\mathbf{k})t} \Psi_+(\mathbf{k}) d\mathbf{k} = (-it)^{-N} \int e^{-i\phi_+(\mathbf{k})t} (D^*)^N \Psi_+(\mathbf{k}) d\mathbf{k},$$

where $D^* \Psi(\mathbf{k}) = \nabla \cdot [\nabla \bar{\phi}(\mathbf{k}) \Psi(\mathbf{k}) / |\nabla \phi(\mathbf{k})|^2]$ is the adjoint operator to D . This implies the bound

$$(5.20) \quad |\psi_{\pm}(t, \mathbf{v}t)| \leq C_N |t|^{-N}, \quad t > 0$$

for $\mathbf{v} \notin \mathbf{V}_a(\mathbf{v}_+)$. The bound is close to (5.11) for $\mathbf{x} = \mathbf{v}t$, however the dependence on \mathbf{x} in the right hand side has to be clarified.

To do this, we analyze with details the operator $(D^*)^N$:

$$(5.21) \quad (D^*)^N \Psi_+(\mathbf{k}) = \sum_{|\alpha| \leq N} C_{N,\alpha}(\mathbf{k}) \partial^\alpha \Psi_+(\mathbf{k}),$$

where the coefficients $C_{N,\alpha}$ admit the bound

$$(5.22) \quad |C_{N,\alpha}(\mathbf{k})| \leq \frac{C_N}{|\nabla \phi(\mathbf{k})|^N}, \quad \mathbf{k} \in \text{supp } \Psi_+,$$

with a constant C_N which does not depend on $\mathbf{v} \notin \mathbf{V}_a(\mathbf{v}_+)$. Therefore, (5.18) implies that

$$(5.23) \quad |t^{-N} C_{N,\alpha}(\mathbf{k})| \leq \frac{C}{t^N |\nabla \phi(\mathbf{k})|^N} \leq \frac{C(\varepsilon)}{(|t| + |\mathbf{v}t|)^N} = \frac{C(\varepsilon)}{(|t| + |\mathbf{x}|)^N}, \quad \frac{\mathbf{x}}{t} \notin \mathbf{V}_a(\mathbf{v}_+), \quad \mathbf{k} \in \text{supp } \Psi_+.$$

Substituting (5.21) and (5.23) in (5.19) we get (5.11).

ii) It remains to consider $\mathbf{v} \in \mathbf{V}_r(\mathbf{k}_*)$. This means that $\mathbf{v} = \nabla \omega(\mathbf{k})$ with a point $\mathbf{k} \in B_r(\mathbf{k}_*)$ which is a solution to the system (5.14). Then the integral (5.13) is called the *Fresnel* integral and its asymptotics is $\sim |t|^{-3/2}$ (see [32]). ■

Remark 5.5 *The asymptotics (5.11), (5.12) means that the energy of the field $\psi_{\pm}(t, \mathbf{x})$ outside a ball $|\mathbf{x} - \mathbf{v}_{\pm}t| \leq a|t|$ decays rapidly, while inside it is about constant since the energy is a quadratic form. Therefore, the wave packet $\psi_{\pm}(t, \mathbf{x})$ of the free Klein-Gordon equation (5.1) moves like a free particle of the size $\sim a|t|$ and with the group velocity*

$$\mathbf{v}_{\pm} = \pm \nabla \omega(\mathbf{k}_*) = \frac{\mathbf{k}_*}{\sqrt{1 + \mathbf{k}_*^2}}.$$

This formula means that the velocity \mathbf{v}_{\pm} corresponds to the relativistic momentum $\pm \mathbf{k}_$ of the particle.*

Exercise 5.6 Analyze the wave packet propagation for the **free** Schrödinger equation

$$(5.24) \quad \left\{ \begin{array}{l} -i\dot{\psi}(t, \mathbf{x}) = \frac{1}{2\mu} \Delta\psi(t, \mathbf{x}) \\ \psi(0, \mathbf{x}) = \psi_0(\mathbf{x}), \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3.$$

Prove that the packets move like **free non-relativistic particles** of the size $\mathcal{O}(r|t|/|\mathbf{k}_*|)$ with mass μ and momentum \mathbf{k}_* .

Hint: $\omega = \mathbf{k}^2/2\mu$, hence the **group velocity** \mathbf{v} equals $\nabla\omega = \mathbf{k}/\mu$.

5.2 WKB Asymptotics for Schrödinger Equation with a Maxwell Field

Let us write the Lorentz equation (4.21) in Hamilton form with the Hamiltonian (4.32):

$$(5.25) \quad \left\{ \begin{array}{l} \dot{\mathbf{x}} = H_{\mathbf{p}}(\mathbf{x}, \mathbf{p}, t) = \frac{1}{\mu}(p - \frac{e}{c}A(t, \mathbf{x})), \\ \dot{\mathbf{p}} = -H_{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t) = -e\phi_{\mathbf{x}}(t, \mathbf{x}) + \frac{e}{\mu c}A_{\mathbf{x}}(t, \mathbf{x}), \end{array} \right.$$

where e is the charge of the particle and μ its mass. E.Schrödinger associated with the Hamilton system the wave equation (2.2):

$$(5.26) \quad (i\hbar\partial_t - e\phi(t, \mathbf{x}))\psi(t, \mathbf{x}) = \frac{1}{2\mu}(-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}A(t, \mathbf{x}))^2\psi(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Let us demonstrate that the short-wave solutions to (5.26) are governed by the Hamilton equations (5.25). More precisely, let us consider the Cauchy problem for (5.26) with the initial condition

$$(5.27) \quad \psi|_{t=0} = a_0(\mathbf{x})e^{iS_0(\mathbf{x})/\hbar}, \quad \mathbf{x} \in \mathbb{R}^3,$$

where $S_0(\mathbf{x})$ is a real function. Let us denote by $(\mathbf{x}(t, \mathbf{x}_0), \mathbf{p}(t, \mathbf{x}_0))$ the solution to the Hamilton equations (5.25) with the initial conditions (38.8):

$$(5.28) \quad \mathbf{x}|_{t=0} = \mathbf{x}_0, \quad \mathbf{p}|_{t=0} = \nabla S_0(\mathbf{x}_0).$$

The solution exists for $|t| < T(\mathbf{x}_0)$ where $T(\mathbf{x}_0) > 0$.

Definition 5.7 The curve $x = \mathbf{x}(t, \mathbf{x}_0)$ is the **ray** of the Cauchy problem (5.26), (5.27) starting at the point x_0 .

Definition 5.8 The **ray tube** or **ray beam** emanating from the initial function (5.27) is the set $\mathcal{T} = \{(t, \mathbf{x}(t, \mathbf{x}_0)) \in \mathbb{R}^4 : |t| < T(\mathbf{x}_0), \mathbf{x}_0 \in \text{supp } a_0\}$.

The following theorem means roughly speaking that for $\hbar \ll 1$ the set \mathcal{T} is the support of the wave function $\psi(t, \mathbf{x})$, the solution to the Cauchy problem (5.26), (5.27). Let us assume that the potentials $\phi(t, \mathbf{x}), A(t, \mathbf{x}) \in C^\infty(\mathbb{R}^4)$ and $a_0, S_0 \in C^\infty(\mathbb{R}^3)$. Then the map $\mathbf{x}_0 \mapsto \mathbf{x}(t, \mathbf{x}_0)$ is a local C^∞ -diffeomorphism of \mathbb{R}^3 for small $|t|$. We will construct the formal *Debye expansion*

$$(5.29) \quad \psi(t, \mathbf{x}) \sim \left(\sum_{k=0}^{\infty} \hbar^k a_k(t, \mathbf{x}) \right) e^{iS(t, \mathbf{x})/\hbar}, \quad \hbar \rightarrow 0.$$

Theorem 5.9 Let the map $\mathbf{x}_0 \mapsto \mathbf{x}(t, \mathbf{x}_0)$ be a diffeomorphism of \mathbb{R}^3 for $|t| < T$. Then the formal expansion (5.29) exists for $|t| < T$ and is identically zero outside \mathcal{T} , i.e.

$$(5.30) \quad a_k(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \notin \mathcal{T}, \quad |t| < T, \quad k = 0, 1, 2, \dots$$

Proof First, let us define the phase function $S(t, \mathbf{x})$ as the solution to the Cauchy problem (38.7) with $N = 3$:

$$(5.31) \quad \begin{cases} -\dot{S}(t, \mathbf{x}) = H(\mathbf{x}, \nabla S(t, \mathbf{x}), t), & \mathbf{x} \in \mathbb{R}^3, \quad |t| < T, \\ S|_{t=0} = S_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^3. \end{cases}$$

The solution exists by Theorem 38.10. Further, let us substitute $\psi(t, \mathbf{x}) = a(t, \mathbf{x})e^{iS(t, \mathbf{x})/\hbar}$ into the Schrödinger equation (5.26). Then Equation (5.31) implies the following *transport equation* for the amplitude $a(t, \mathbf{x})$:

$$(5.32) \quad \begin{cases} \dot{a}(t, \mathbf{x}) = -\frac{1}{m}[\nabla S(t, \mathbf{x}) - \frac{e}{c}A(t, \mathbf{x})] \cdot \nabla a(t, \mathbf{x}) \\ \quad + \frac{1}{2\mu}[\Delta S(t, \mathbf{x}) + \nabla \cdot A(t, \mathbf{x})]a(t, \mathbf{x}) + \frac{i\hbar}{2\mu}\Delta a(t, \mathbf{x}) \\ \quad =: -La(t, \mathbf{x}) + B(t, \mathbf{x})a(t, \mathbf{x}) + \frac{i\hbar}{2\mu}d(t, \mathbf{x}), & |t| < T, \\ a|_{t=0} = a_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^3, \end{cases}$$

where L is the first order differential operator $La(t, \mathbf{x}) := \frac{1}{m}[\nabla S(t, \mathbf{x}) - \frac{e}{c}A(t, \mathbf{x})] \cdot \nabla a(t, \mathbf{x})$, $B(t, \mathbf{x})$ is the function $\frac{1}{2\mu}[\Delta S(t, \mathbf{x}) + \nabla \cdot A(t, \mathbf{x})]$ and $d(t, \mathbf{x}) := \Delta a(t, \mathbf{x})$. Let us note that (5.25) implies

$$(5.33) \quad \dot{a}(t, \mathbf{x}) + La(t, \mathbf{x}) = \frac{d}{dt}a(t, \mathbf{x}(t, \mathbf{x}_0)).$$

Let us express all functions in *ray coordinates* (t, \mathbf{x}_0) : $\tilde{a}(t, \mathbf{x}_0) := a(t, \mathbf{x}(t, \mathbf{x}_0))$, $\tilde{B}(t, \mathbf{x}_0) := B(t, \mathbf{x}(t, \mathbf{x}_0))$ etc. Then (5.32) can be rewritten as

$$(5.34) \quad \begin{cases} \frac{d}{dt}\tilde{a}(t, \mathbf{x}_0) = \tilde{B}(t, \mathbf{x}_0)\tilde{a}(t, \mathbf{x}_0) + \frac{i\hbar}{2\mu}\tilde{d}(t, \mathbf{x}_0), & |t| < T, \\ \tilde{a}(0, \mathbf{x}_0) = a_0(\mathbf{x}_0), & \mathbf{x}_0 \in \mathbb{R}^3. \end{cases}$$

Let us substitute in the first equation the formal expansion $\tilde{a}(t, \mathbf{x}_0) \sim \sum_{k=0}^{\infty} \hbar^k \tilde{a}_k(t, \mathbf{x}_0)$. Equating formally the terms with identical powers of \hbar , we get the recursive *transport equations*

$$(5.35) \quad \left. \begin{aligned} \frac{d}{dt}\tilde{a}_0(t, \mathbf{x}_0) &= \tilde{B}(t, \mathbf{x}_0)\tilde{a}_0(t, \mathbf{x}_0), \\ \frac{d}{dt}\tilde{a}_1(t, \mathbf{x}_0) &= \tilde{B}(t, \mathbf{x}_0)\tilde{a}_1(t, \mathbf{x}_0) + \frac{i}{2\mu}\tilde{d}_0(t, \mathbf{x}_0), \\ \dots & \\ \frac{d}{dt}\tilde{a}_k(t, \mathbf{x}_0) &= \tilde{B}(t, \mathbf{x}_0)\tilde{a}_k(t, \mathbf{x}_0) + \frac{i}{2\mu}\tilde{d}_{k-1}(t, \mathbf{x}_0), \\ \dots & \end{aligned} \right| \quad |t| < T,$$

where $\tilde{d}_0(t, \mathbf{x}_0)$ is the function $d_0(t, x) := \Delta a_0(t, x)$ expressed in ray coordinates, etc. It remains to substitute the same expansion into the initial conditions (5.34) which gives

$$(5.36) \quad \tilde{a}_0(0, \mathbf{x}_0) = a_0(\mathbf{x}_0), \quad \tilde{a}_1(0, \mathbf{x}_0) = 0, \quad \dots, \quad \tilde{a}_k(0, \mathbf{x}_0) = 0, \dots$$

Now (5.35) and (5.36) imply that $\tilde{a}_0(t, \mathbf{x}_0) = 0$, $|t| < T$, if $\mathbf{x}_0 \notin \text{supp } a_0$. Hence, also $\tilde{d}_0(t, \mathbf{x}_0) = 0$, $|t| < T$, if $\mathbf{x}_0 \notin \text{supp } a_0$. Therefore, (5.35) and (5.36) imply that $\tilde{a}_1(t, \mathbf{x}_0) = 0$, $|t| < T$, if $\mathbf{x}_0 \notin \text{supp } a_0$, etc. ■

Exercise 5.10 *Prove the transport equation (5.32). **Hint:** Substitute $\psi = ae^{iS/\hbar}$ into the Schrödinger equation (5.26) and divide by $e^{iS/\hbar}$. Then Equation (5.31) formally follows by setting $\hbar = 0$, and afterwards, (5.32) also follows.*

6 Schrödinger Equation and Heisenberg Representation

According to Schrödinger, the quantum mechanical electron in an external electromagnetic field is described by the corresponding Schrödinger equation for the electron field. The results of the last section then imply that, in the classical limit, a ray of electrons propagates according to the Lorentz dynamics. The symmetries of the Schrödinger equation lead to conserved physical observables of the electron field, like energy, momentum, angular momentum, and electric charge. All observables are quadratic forms of linear operators in the Hilbert space of electron wave functions, therefore the quantum observables are identified with linear operators in this Hilbert space.

6.1 Electrons and Cathode Rays

The *electron* was discovered by J.J.Thomson around 1897 in *cathode rays*. He systematized observations of the deflection of rays in Maxwell fields. The observations demonstrated that the deflections may be described by the Lorentz dynamics (4.21) or (5.25), with a fixed ratio e/μ close to its present value. Later Kauffmann [58] confirmed these observations with high accuracy, $e/\mu \approx -1.76 \cdot 10^8 \text{C/g}$. The charge of the electron can be evaluated from Faraday's electrolytic law to be $e \approx -96500 \text{C}/6.06 \cdot 10^{23} \approx 1.6 \cdot 10^{-19} \text{C} \approx -4.77 \cdot 10^{-10} \text{esu}$. In 1913 Millikan confirmed this result, $e \approx 1.60 \cdot 10^{-19} \text{C} \approx -4.77 \cdot 10^{-10} \text{esu}$, leading to an electron mass of $\mu \approx 9.1 \cdot 10^{-28} \text{g}$. Therefore, by Theorem 5.9, the *short-wave* cathode rays can also be described by the Schrödinger equation (5.26) with the electron mass μ and the *negative* electron charge $e < 0$:

$$(6.1) \quad [i\hbar\partial_t - e\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Remark 6.1 Let us note that the charge of the cathode rays is not fixed, hence the equation (6.1) describes the *short-wave* cathode rays of an arbitrary charge. It means that the Schrödinger equation describes rather the dynamics of the *electron field*, with a fixed ratio e/μ , than the dynamics of a particle with the charge e .

Remark 6.2 Theorem 5.9 means that the Schrödinger equation (6.1), with any small $\hbar \ll 1$, agrees with the classical Lorentz equations (4.21) or (5.25). The actual value of the constant is fixed by the Planck relation $\hbar = a/k \approx 1.05 \cdot 10^{-27} \text{erg} \cdot \text{sec}$. Here a is the parameter in the Wien experimental formula (1.4) and k is the Boltzmann constant. This identification of a small parameter \hbar in the Schrödinger equation follows by the experimental and theoretical development from the Kirchhoff spectral law to de Broglie's wave-particle duality.

6.2 Quantum Stationary States

Let us consider the case of a *static* Maxwell field with the potentials $\phi(t, \mathbf{x}) \equiv \phi(\mathbf{x})$ and $\mathbf{A}(t, \mathbf{x}) \equiv \mathbf{A}(\mathbf{x})$. The corresponding Schrödinger equation (6.1) becomes,

$$(6.2) \quad i\hbar\partial_t\psi(t, \mathbf{x}) = \mathcal{H}\psi(t, \mathbf{x}) := \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]^2\psi(t, \mathbf{x}) + e\phi(\mathbf{x})\psi(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Then the energy is conserved (see (3.1)):

$$(6.3) \quad \int_{\mathbb{R}^3} \left(\frac{1}{2\mu} |[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]\psi(t, \mathbf{x})|^2 + e\phi(\mathbf{x})|\psi(t, \mathbf{x})|^2 \right) d\mathbf{x} = E, \quad t \in \mathbb{R}.$$

Definition 6.3 Quantum Stationary States for Equation (6.2) are nonzero finite energy solutions of the type

$$(6.4) \quad \psi(t, \mathbf{x}) = \psi_\omega(\mathbf{x})e^{-i\omega t}$$

with an $\omega \in \mathbb{R}$.

Substituting (6.4) into (6.2), we get the *stationary Schrödinger equation*

$$(6.5) \quad \hbar\omega\psi_\omega(\mathbf{x}) = \mathcal{H}\psi_\omega(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3,$$

which constitutes an *eigenvalue problem*. Substituting (6.4) into (6.3) and using (6.5), we get for the energy

$$(6.6) \quad E = E_\omega = \hbar\omega \int_{\mathbb{R}^3} |\psi_\omega(\mathbf{x})|^2 d\mathbf{x} < \infty.$$

Hence, the quantum stationary states correspond to the eigenfunctions ψ_ω from the Hilbert space $\mathcal{E} := L^2(\mathbb{R}^3)$. Let us assume the standard **normalization condition** (see Remark 7.1)

$$(6.7) \quad \int_{\mathbb{R}^3} |\psi_\omega(\mathbf{x})|^2 d\mathbf{x} = 1.$$

Then (6.6) becomes (cf. (P) from the Introduction)

$$(6.8) \quad E_\omega = \hbar\omega,$$

and (6.5) takes the form

$$(6.9) \quad E_\omega\psi_\omega(\mathbf{x}) = \mathcal{H}\psi_\omega(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

6.3 Four Conservation Laws for Schrödinger Equation

Let us state the four classical conservation laws for the Schrödinger equation (6.2). We adjust the conserved quantities by some factors.

I. Energy Conservation: The total energy is conserved if the Maxwell potentials do not depend on time t , i.e. $\phi(t, \mathbf{x}) = \phi(\mathbf{x})$, $\mathbf{A}(t, \mathbf{x}) = \mathbf{A}(\mathbf{x})$:

$$(6.10) \quad E(t) := \int \left(\frac{1}{2\mu} |[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(\mathbf{x})]\psi(t, \mathbf{x})|^2 + e\phi(\mathbf{x})|\psi(t, \mathbf{x})|^2 \right) d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}.$$

This follows from Theorem 3.4 by (3.1) and (40.17) .

II. Momentum Conservation The component \mathbf{p}_n , $n = 1, 2, 3$, of the total momentum, is conserved if the Maxwell potentials do not depend on \mathbf{x}_n :

$$(6.11) \quad \mathbf{p}_n(t) := -i\hbar \int \left[\psi(t, \mathbf{x}) \cdot \nabla_n \psi(t, \mathbf{x}) \right] d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}.$$

This follows from Theorem 3.8 by (3.16) and (40.19).

III. Angular Momentum Conservation Let us assume that the Maxwell fields are axially symmetric with respect to rotations around the axis \mathbf{x}_n , i.e.

$$(6.12) \quad \phi(t, R_n(s)\mathbf{x}) \equiv \phi(t, \mathbf{x}) \quad \text{and} \quad \mathbf{A}(t, R_n(s)\mathbf{x}) \equiv R_n^t(s)\mathbf{A}(t, \mathbf{x}), \quad s \in \mathbb{R}.$$

Then corresponding component of the total angular momentum is conserved:

$$(6.13) \quad \mathbf{L}_n(t) := -i\hbar \int \left[\psi(t, \mathbf{x}) \cdot (\mathbf{x} \times \nabla)_n \psi(t, \mathbf{x}) \right] d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}.$$

This follows from Theorem 3.13 by (3.23) and (40.21) since the Lagrangian density (2.6) satisfies the rotation-invariance condition (3.21) by (6.12).

IV. Charge Conservation The total charge is conserved:

$$(6.14) \quad Q(t) := e \int |\psi(t, \mathbf{x})|^2 d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}.$$

This follows from Theorem 3.18 by (3.25) and (40.21).

Exercise 6.4 Check the rotation-invariance condition (3.21) for the Maxwell potentials satisfying (6.12).

Exercise 6.5 Check that the axial symmetry condition (6.12) is equivalent to the corresponding axial symmetry of the Maxwell field: for every $s \in \mathbb{R}$,

$$(6.15) \quad \mathbf{E}(t, R_n(s)\mathbf{x}) \equiv R_n(s)\mathbf{E}(t, \mathbf{x}), \quad \mathbf{B}(t, R_n(s)\mathbf{x}) \equiv R_n(s)\mathbf{B}(t, \mathbf{x}).$$

6.4 Quantum Observables and Heisenberg Representation

The charge conservation (6.14) implies that the *complex* Hilbert space $\mathcal{E} := L^2(\mathbb{R}^3)$ is invariant with respect to the dynamics defined by the Schrödinger equation. For $\psi_{1,2} \in \mathcal{E}$, let us denote by $\langle \psi_1, \psi_2 \rangle$ the Hermitian scalar product:

$$(6.16) \quad \langle \psi_1, \psi_2 \rangle := \int \psi_1(\mathbf{x}) \overline{\psi_2(\mathbf{x})} d\mathbf{x}.$$

All conserved quantities $E, \mathbf{p}_n, \mathbf{L}_n, Q$ are quadratic forms in the phase space \mathcal{E} :

$$(6.17) \quad E = \langle \psi, \hat{E}\psi \rangle, \quad \mathbf{p}_n = \langle \psi, \hat{\mathbf{p}}_n\psi \rangle, \quad \mathbf{L}_n = \langle \psi, \hat{\mathbf{L}}_n\psi \rangle, \quad Q = \langle \psi, \hat{Q}\psi \rangle.$$

Here $\hat{E}, \hat{\mathbf{p}}_n, \hat{\mathbf{L}}_n, \hat{Q}$ stand for the corresponding operators of the quadratic forms:

$$(6.18) \quad \hat{E} = \mathcal{H}, \quad \hat{\mathbf{p}}_n = -i\hbar\nabla_n, \quad \hat{\mathbf{L}}_n = -i\hbar(\mathbf{x} \times \nabla)_n, \quad \hat{Q} = I.$$

This motivates the following

Definition 6.6 *i)* A **quantum observable** is a linear operator $\hat{\mathbf{M}}$ in the Hilbert space \mathcal{E} .
ii) The quadratic form $\mathbf{M}(\psi) := \langle \psi, \hat{\mathbf{M}}\psi \rangle$ is a **mean value** of the observable at the state $\psi \in \mathcal{E}$.

Example 6.7 The multiplication operator $\hat{\mathbf{x}}_k\psi = \mathbf{x}_k\psi(\mathbf{x})$ is the quantum observable of the k -th **coordinate**.

Example 6.8 The (squared) absolute value of the total angular momentum is defined by $\hat{\mathbf{L}}^2 := \hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2$.

Exercise 6.9 Prove that all mean values $\langle \psi, \hat{\mathbf{M}}\psi \rangle$ of a quadratic form are real iff the operator $\hat{\mathbf{M}}$ is symmetric.

Exercise 6.10 Check that all the operators $\hat{E}, \hat{\mathbf{p}}_n, \hat{\mathbf{L}}_n$ are symmetric. **Hint:** All the mean values are real since they coincide with the total energy, momentum, etc, defined by the real Lagrangian density (2.6).

Exercise 6.11 Check the formula

$$(6.19) \quad \hat{\mathbf{L}}_n = -i\hbar \frac{\partial}{\partial \varphi_n},$$

where φ_n is the angle of rotation around the vector e_n in a positive direction. **Hint:** Consider $n = 3$ and choose the polar coordinate in the plane $\mathbf{x}_1, \mathbf{x}_2$.

Now let us assume that $\phi(t, \mathbf{x}) = \phi(\mathbf{x})$, $\mathbf{A}(t, \mathbf{x}) = \mathbf{A}(\mathbf{x})$ and denote by $U(t)$, $t \in \mathbb{R}$, the dynamical group of the Schrödinger equation in the Hilbert space \mathcal{E} : by definition, $U(t)\psi(0, \cdot) = \psi(t, \cdot)$ for any solution $\psi(t, \mathbf{x})$ to the equation. The Schrodinger equation implies that

$$(6.20) \quad \dot{U}(t) = -\frac{i}{\hbar} \mathcal{H}U(t) = -\frac{i}{\hbar} U(t)\mathcal{H}.$$

Exercise 6.12 Prove (6.20).

The charge conservation (6.14) means that $U(t)$ is a unitary operator for all $t \in \mathbb{R}$.

Let us note that

$$(6.21) \quad \hat{E} = \mathcal{H} = \frac{1}{2\mu} \left[\hat{\mathbf{p}} - \frac{e}{c} \mathbf{A}(\hat{\mathbf{x}}) \right]^2 + e\phi(\hat{\mathbf{x}}), \quad \hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}, \quad [\hat{\mathbf{x}}_k, \hat{\mathbf{p}}_n] = i\hbar \delta_{kn},$$

where $[\cdot, \cdot]$ stands for the commutator $[A, B] := AB - BA$.

Exercise 6.13 Check the identities (6.21). **Hint:** For polynomial potentials $\mathbf{A}(\mathbf{x})$, $\phi(\mathbf{x})$, the operator of multiplication by $\mathbf{A}(\mathbf{x})$ resp. $\phi(\mathbf{x})$ coincides with $\mathbf{A}(\hat{\mathbf{x}})$ resp. $\phi(\hat{\mathbf{x}})$.

Definition 6.14 The Heisenberg representation of a quantum observable $\hat{\mathbf{M}}$ is an operator function $\hat{\mathbf{M}}(t) := U(-t)\hat{\mathbf{M}}U(t)$, $t \in \mathbb{R}$.

The definition implies that

$$(6.22) \quad \mathbf{M}(\psi(t)) = \langle \psi(0), \hat{\mathbf{M}}(t)\psi(0) \rangle, \quad t \in \mathbb{R}$$

for any solution $\psi(t) := \psi(t, \cdot)$ to the the Schrödinger equation.

Exercise 6.15 Check (6.22).

Exercise 6.16 Check the identities

$$(6.23) \quad \hat{E}(t) = \mathcal{H} = \frac{1}{2\mu} \left[\hat{\mathbf{p}}(t) - \frac{e}{c} \mathbf{A}(\hat{\mathbf{x}}(t)) \right]^2 + e\phi(\hat{\mathbf{x}}(t)), \quad \hat{\mathbf{L}}(t) = \hat{\mathbf{x}}(t) \times \hat{\mathbf{p}}(t), \quad [\hat{\mathbf{x}}_k(t), \hat{\mathbf{p}}_n(t)] = i\hbar \delta_{kn}.$$

Hint: First consider polynomial potentials $\mathbf{A}(\mathbf{x})$, $\phi(\mathbf{x})$ and use (6.21) together with the commutation $\mathcal{H}U(t) = U(t)\mathcal{H}$ from (6.20).

Exercise 6.17 Check the identities

$$(6.24) \quad [\hat{\mathbf{x}}_k(t), \hat{\mathbf{p}}_n^N(t)] = i\hbar \delta_{kn} N \hat{\mathbf{p}}_n^{N-1}, \quad [\mathbf{A}(\hat{\mathbf{x}}(t)), \hat{\mathbf{p}}_n(t)] = i\hbar \delta_{kn} \mathbf{A}_{\mathbf{x}_n}(\hat{\mathbf{x}}(t)), \quad [\phi(\hat{\mathbf{x}}(t)), \hat{\mathbf{p}}_n(t)] = i\hbar \delta_{kn} \phi_{\mathbf{x}_n}(\hat{\mathbf{x}}(t)),$$

for any $N = 1, 2, \dots$ **Hint:** First consider polynomial potentials $\mathbf{A}(\mathbf{x})$, $\phi(\mathbf{x})$ and apply the last formula of (6.23) by induction.

Let us derive a dynamical equation for $\hat{\mathbf{M}}(t)$: (6.20) implies the **Heisenberg equations**

$$(6.25) \quad \dot{\hat{\mathbf{M}}}(t) = \frac{i}{\hbar}[\mathcal{H}, \hat{\mathbf{M}}(t)] = \frac{i}{\hbar}[\hat{E}(t), \hat{\mathbf{M}}(t)], \quad t \in \mathbb{R}.$$

Exercise 6.18 Check (6.25).

Lemma 6.19 The mean value, $\mathbf{M}(\psi(t))$, is conserved if $[\mathcal{H}, \hat{\mathbf{M}}] = 0$.

Proof By (6.2),

$$(6.26) \quad \begin{aligned} \frac{d}{dt}\mathbf{M}(\psi(t)) &= \langle \dot{\psi}(t), \hat{\mathbf{M}}\psi(t) \rangle + \langle \psi(t), \hat{\mathbf{M}}\dot{\psi}(t) \rangle = -\langle \frac{i}{\hbar}\mathcal{H}\psi(t), \hat{\mathbf{M}}\psi(t) \rangle - \langle \psi(t), \hat{\mathbf{M}}\frac{i}{\hbar}\mathcal{H}\psi(t) \rangle \\ &= -\frac{i}{\hbar}[\langle \mathcal{H}\psi(t), \hat{\mathbf{M}}\psi(t) \rangle - \langle \psi(t), \hat{\mathbf{M}}\mathcal{H}\psi(t) \rangle] = -\frac{i}{\hbar}\langle \psi(t), [\mathcal{H}, \hat{\mathbf{M}}]\psi(t) \rangle, \quad t \in \mathbb{R}. \quad \blacksquare \end{aligned}$$

Exercise 6.20 Prove the angular momentum conservation (6.13), by using Lemma 6.19. **Hint:** Check the commutation $[\mathcal{H}, \hat{\mathbf{L}}_n] = 0$ under condition (6.12).

Let us write the Heisenberg equations for the observables $\hat{\mathbf{x}}(t)$ and $\hat{\mathbf{p}}(t)$. Then we obtain the system

$$(6.27) \quad \begin{cases} \dot{\hat{\mathbf{x}}}(t) &= \frac{1}{\mu}(\hat{\mathbf{p}}(t) - \frac{e}{c}A(t, \hat{\mathbf{x}}(t))), \\ \dot{\hat{\mathbf{p}}}(t) &= -e\phi_{\mathbf{x}}(t, \hat{\mathbf{x}}(t)) + \frac{e}{\mu c}A_{\mathbf{x}}(t, \hat{\mathbf{x}}(t)). \end{cases}$$

Exercise 6.21 Check (6.27). **Hint:** Use (6.25) and (6.24) with $N = 1, 2$.

Let us note that Equations (6.27) formally coincide with the *classical* Hamilton system (5.25).

7 Coupling to the Maxwell Equations

The simultaneous evolution of the full system of electron wave function and Maxwell field is determined by the coupled Schrödinger and Maxwell equations. Stationary states of this coupled system would provide natural candidates for physical states and transitions, but the existence of these stationary states as well as the transition of general solutions to these stationary states for large times is an open problem. Approximate, iterative solutions to the coupled system are provided by the Born approximation.

7.1 Charge and Current Densities and Gauge Invariance

Let us determine the dynamics of the Maxwell field in presence of the wave field ψ governed by the Schrödinger equation. The Lagrangian density of the free Maxwell field is known from (4.18), so we have to modify only the interaction term in (4.18). The interaction term gives the dynamics of the Maxwell field in the presence of given charge-current densities (4.18), so it remains to express the densities in terms of the wave field ψ .

On the other hand, we have shown that the interaction term in (4.18) uniquely determines the Lorentz dynamics (4.21). The Schrödinger equation (6.1) substitutes the Lorentz dynamics (4.21). This suggests to identify the Maxwell-Schrödinger interaction with the interaction term from the Lagrangian density \mathcal{L}_S of the Schrödinger equation (6.1). That is, we add the Lagrangian density \mathcal{L}_f of the free Maxwell field to the Lagrangian density \mathcal{L}_S of the Schrödinger equation (6.1) and get the Lagrangian density \mathcal{L}_{MS} of the coupled Maxwell-Schrödinger equations,

$$(7.1) \quad \mathcal{L}_{MS} = [i\hbar\partial_t - e\phi]\psi \cdot \psi - \frac{1}{2\mu} \sum_{k=1}^3 |[-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k]\psi|^2 - \frac{1}{16\pi} \mathcal{F}^{\alpha\beta} \mathcal{F}_{\alpha\beta},$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$, $\mathcal{A}^\nu = (\phi, \mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3)$. The corresponding Euler-Lagrange equations read

$$(7.2) \quad \left\{ \begin{array}{l} [i\hbar\partial_t - e\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2 \psi(t, \mathbf{x}), \\ \frac{1}{4\pi} \nabla_\alpha \mathcal{F}^{\alpha\beta}(t, \mathbf{x}) = \begin{pmatrix} \rho := e|\psi(t, \mathbf{x})|^2, & \beta = 0 \\ \frac{\mathbf{j}_\beta}{c} := \frac{e}{\mu c} [-i\hbar\nabla_\beta - \frac{e}{c}\mathbf{A}_\beta(t, \mathbf{x})]\psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}), & \beta = 1, 2, 3 \end{pmatrix} \end{array} \right.$$

The system (7.2) describes the dynamics of the wave field ψ in its “own”, induced Maxwell field $\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x})$ generated by the charges and currents of the wave field.

Remark 7.1 *i) The charge-current densities from (7.2) coincide, up to a factor, with the Noether currents (40.23) of the group of internal rotations (39.12).*

ii) The factor is chosen in accordance to the normalization condition (6.7). Namely the condition means that the total charge of the corresponding quantum stationary state is equal to e .

Now let us introduce external potentials $\phi^{\text{ext}}(t, \mathbf{x}), \mathbf{A}^{\text{ext}}(t, \mathbf{x})$ of the Maxwell field generated by some external sources. We formalize the introduction through the Lagrangian density

$$(7.3) \quad \mathcal{L}_{MS} = [i\hbar\partial_t - e(\phi + \phi^{\text{ext}})]\psi \cdot \psi - \frac{1}{2\mu} \sum_{k=1}^3 |[-i\hbar\nabla_k - \frac{e}{c}(\mathbf{A}_k + \mathbf{A}_k^{\text{ext}})]\psi|^2 - \frac{1}{16\pi} \mathcal{F}^{\alpha\beta} \mathcal{F}_{\alpha\beta}.$$

The corresponding equations read

$$(7.4) \quad \left\{ \begin{array}{l} [i\hbar\partial_t - e(\phi(t, \mathbf{x}) + \phi^{\text{ext}}(t, \mathbf{x}))]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}(\mathbf{A}(t, \mathbf{x}) + \mathbf{A}^{\text{ext}}(t, \mathbf{x}))]^2\psi(t, \mathbf{x}), \\ \frac{1}{4\pi} \nabla_{\alpha} \mathcal{F}^{\alpha\beta}(t, \mathbf{x}) = \left(\begin{array}{l} \rho := e|\psi(t, \mathbf{x})|^2 \\ \frac{\mathbf{j}_{\beta}}{c} := \frac{e}{\mu c} [-i\hbar\nabla_{\beta} - \frac{e}{c}(\mathbf{A}_{\beta}(t, \mathbf{x}) + \mathbf{A}_{\beta}^{\text{ext}}(t, \mathbf{x}))]\psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}) \end{array} \right) \end{array} \right. .$$

Remark 7.2 The tensor $\mathcal{F}^{\alpha\beta}$ is defined by the potentials $\phi(t, \mathbf{x})$, $\mathbf{A}(t, \mathbf{x})$, hence the system admits the solution $\psi(t, \mathbf{x}) = 0$, $\phi(t, \mathbf{x}) = 0$, $\mathbf{A}(t, \mathbf{x}) = 0$ corresponding to the absence of matter.

The gauge transformation (4.4) does not change the Maxwell fields $E(t, \mathbf{x})$, $B(t, \mathbf{x})$ for any function $\chi(t, \mathbf{x}) \in C^1(\mathbb{R}^4)$. Therefore, it would be natural to expect that the solutions to the coupled equations (7.2), (7.4) also do not change too much under this transformation. Indeed, we can complete the transformation of the potentials (4.4) with a corresponding transformation of the wave function:

$$(7.5) \quad \left\{ \begin{array}{l} \phi(t, \mathbf{x}) \mapsto \phi(t, \mathbf{x}) + \frac{1}{c} \dot{\chi}(t, \mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) \mapsto \mathbf{A}(t, \mathbf{x}) - \nabla_{\mathbf{x}}\chi(t, \mathbf{x}), \\ \psi(t, \mathbf{x}) \mapsto e^{-i\frac{e}{\hbar c}\chi(t, \mathbf{x})} \psi(t, \mathbf{x}). \end{array} \right.$$

It is easy to check that the new functions also provide a solution to Equations (7.2) resp. (7.4). Moreover, the transformations (7.5) do not change the electric charge and current densities $\rho(t, \mathbf{x})$ and $\mathbf{j}(t, \mathbf{x})$ in (7.2) and (7.4).

7.2 Electron Beams and Heisenberg's Uncertainty Principle

Plane wave as an electron beam

According to the de Broglie's wave-particle duality (1.7), a free electron beam is described by a plane wave

$$(7.6) \quad \psi(t, \mathbf{x}) = C e^{i(\mathbf{k}\mathbf{x} - \omega t)}, \quad \mathbf{k} \neq 0$$

which is a solution to the free Schrödinger equation, i.e., without an external Maxwell field. Hence,

$$(7.7) \quad \hbar\omega = \frac{\hbar^2}{2\mu} \mathbf{k}^2 > 0.$$

By (7.2), the corresponding charge and electric current densities are given by

$$(7.8) \quad \rho = e|\psi(t, \mathbf{x})|^2 = e|C|^2,$$

$$(7.9) \quad \mathbf{j} = \frac{e}{\mu} [-i\hbar\nabla\psi(t, \mathbf{x})] \cdot \psi(t, \mathbf{x}) = \frac{e\hbar\mathbf{k}}{\mu} |C|^2.$$

Respectively, the electron density, velocity and momentum are defined by

$$(7.10) \quad \left\{ \begin{array}{l} d := \rho/e = |C|^2, \\ \mathbf{v} = \mathbf{j}/\rho = \frac{\hbar\mathbf{k}}{\mu}, \\ \mathbf{p} = \mu\mathbf{v} = \hbar\mathbf{k}. \end{array} \right.$$

Then the energy density (see (6.10)) admits the following representation:

$$(7.11) \quad e = \frac{1}{2\mu} |[-i\hbar\nabla]\psi(t, \mathbf{x})|^2 = \frac{\hbar^2}{2\mu} |\mathbf{k}|^2 |C|^2 = \frac{\mu \mathbf{v}^2}{2} d.$$

Therefore, the energy of *one electron* is given by

$$(7.12) \quad E = \frac{e}{d} = \frac{\mu \mathbf{v}^2}{2} = \frac{\hbar^2}{2\mu} |\mathbf{k}|^2 = \hbar\omega$$

according to (7.7).

Heisenberg's Uncertainty Principle

The interpretation of the plane wave as a beam of free electrons cannot be completed by an assignment of coordinates to the electrons. This assignment is possible only asymptotically, for small \hbar . Corresponding bounds were discovered by Heisenberg [44, 75]:

$$(7.13) \quad \Delta \mathbf{x} \Delta \mathbf{p} \sim \hbar,$$

where $\Delta x := \langle \psi, (\hat{\mathbf{x}} - \mathbf{x})^2 \psi \rangle^{1/2}$ resp. $\Delta \mathbf{p} := \langle \psi, (\hat{\mathbf{p}} - \mathbf{p})^2 \psi \rangle^{1/2}$ is an **uncertainty** of an electron coordinate resp. momentum (see (6.17) and (6.18) for the notations). For the plane wave the momentum is known exactly, hence $\Delta \mathbf{p} = 0$. Now (7.13) implies that $\Delta \mathbf{x} = \infty$, in other words, the coordinate is not well-defined.

7.3 Quantum Stationary States and Attractors

Let us consider the system (7.4) corresponding to a *static* external Maxwell field $\phi^{\text{ext}}(\mathbf{x})$, $\mathbf{A}^{\text{ext}}(\mathbf{x})$:

$$(7.14) \quad \left\{ \begin{array}{l} [i\hbar\partial_t - e(\phi(t, \mathbf{x}) + \phi^{\text{ext}}(\mathbf{x}))]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}(\mathbf{A}(t, \mathbf{x}) + \mathbf{A}^{\text{ext}}(\mathbf{x}))]^2 \psi(t, \mathbf{x}), \\ \frac{1}{4\pi} \nabla_{\alpha} \mathcal{F}^{\alpha\beta}(t, \mathbf{x}) = \left(\begin{array}{l} \rho := e|\psi(t, \mathbf{x})|^2 \\ \frac{\mathbf{j}_{\beta}}{c} := \frac{e}{\mu c} [-i\hbar\nabla_{\beta} - \frac{e}{c}(\mathbf{A}_{\beta}(t, \mathbf{x}) + \mathbf{A}_{\beta}^{\text{ext}}(\mathbf{x}))]\psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}) \end{array} \right) . \end{array} \right.$$

Let us generalize the definition of quantum stationary states for Eq. (7.14).

Definition 7.3 ([40]) *i) solitary waves for Eq. (7.14) are nonzero finite energy solutions of the type*

$$(7.15) \quad \psi(t, \mathbf{x}) = \psi_{\omega}(\mathbf{x})e^{-i\omega t}, \quad \phi(t, \mathbf{x}) = \phi_{\omega}(\mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) = \mathbf{A}_{\omega}(\mathbf{x})$$

with an $\omega \in \mathbb{R}$.

*ii) The set $\Omega := \{\omega \in \mathbb{R} : \text{there exist a solitary wave with } (\psi_{\omega}(\mathbf{x}), \phi_{\omega}(\mathbf{x}), \mathbf{A}_{\omega}(\mathbf{x})) \neq 0\}$ is the set of the **solitary eigenfrequencies** and $\mathcal{A} := \{(\psi_{\omega}(\mathbf{x}), \phi_{\omega}(\mathbf{x}), \mathbf{A}_{\omega}(\mathbf{x})) : \omega \in \Omega\}$ is the set of the **solitary amplitudes**.*

Substituting (7.15) into (7.14), we get the corresponding stationary *nonlinear eigenfunction problem*: in the Lorentz gauge we get, as in (4.17),

$$(7.16) \quad \begin{cases} [\omega\hbar - e(\phi_\omega(\mathbf{x}) + \phi^{\text{ext}}(\mathbf{x}))]\psi_\omega(\mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}(\mathbf{A}_\omega(\mathbf{x}) + \mathbf{A}^{\text{ext}}(\mathbf{x}))]^2\psi_\omega(\mathbf{x}), \\ -\frac{1}{4\pi}\Delta\phi_\omega(\mathbf{x}) = \rho := e|\psi_\omega(\mathbf{x})|^2, \\ -\frac{1}{4\pi}\Delta\mathbf{A}_\omega(\mathbf{x}) = \frac{\mathbf{j}_\omega}{c} := \frac{e}{\mu c}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}(\mathbf{A}_\omega(\mathbf{x}) + \mathbf{A}^{\text{ext}}(\mathbf{x}))]\psi_\omega(\mathbf{x}) \cdot \psi_\omega(\mathbf{x}). \end{cases}$$

The problem is nonlinear in contrast to (6.5). For the coupled nonlinear Dirac-Maxwell equations with zero external potentials, the existence of solitary waves is proved in [31].

Bohr's quantum transitions (1.6) could be interpreted mathematically as the following convergence (or attraction) for every finite energy solution $\psi(x, t)$ to the Schrödinger equation:

$$(7.17) \quad \psi(t, \mathbf{x}) \sim \psi_\pm(\mathbf{x})e^{-i\omega_\pm t}, \quad t \rightarrow \pm\infty,$$

where $\hbar\omega_- = E_m$ and $\hbar\omega_+ = E_n$. This would mean that the set of all quantum stationary states $\{\psi(x)e^{-i\omega t}\}$ is an *attractor* of the dynamical system defined by the Schrödinger equation (6.2). However, this is generally impossible for the linear autonomous equation by the principle of superposition. Namely, a linear combination of two linearly independent stationary states, $\{\psi_\pm(x)e^{-i\omega_\pm t}\}$,

$$(7.18) \quad \psi(t, \mathbf{x}) = C_- \psi_-(\mathbf{x})e^{-i\omega_- t} + C_+ \psi_+(\mathbf{x})e^{-i\omega_+ t},$$

is a solution to the Schrödinger equation, but the asymptotics (7.17) does not hold if $C_1 \neq 0$ and $C_2 \neq 0$.

On the other hand, the asymptotics could hold for solutions to the nonautonomous Schrödinger equation (6.1) with the potentials $\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x}) \sim \sin(\omega_+ - \omega_-)t$. Let us note that just this type of Maxwell field results from transition (7.17) according to the second Bohr postulate (see Introduction).

Comment 7.4 Summing up, we suggest that the transitions (7.17) hold for the *autonomous coupled nonlinear Maxwell-Schrödinger equations* (7.14): for every finite energy solution $(\psi(t, \mathbf{x}), \phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x}))$, the following long-time asymptotics hold (cf. (0.2))

$$(7.19) \quad (\psi(t, \mathbf{x}), \phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x})) \sim (\psi_\pm(\mathbf{x})e^{-i\omega_\pm t}, \phi_\pm(\mathbf{x}), \mathbf{A}_\pm(\mathbf{x})), \quad t \rightarrow \pm\infty,$$

where $\omega_\pm \in \mathbb{R}$ and $\psi_\pm, \phi_\pm(\mathbf{x}), \mathbf{A}_\pm(\mathbf{x})$ are the solutions to the stationary system (7.16) with $\omega = \omega_\pm$. This means that the set of all solitary amplitudes \mathcal{A} is the *point attractor* of the coupled equations.

Let us note that an *analog* of these transitions was proved recently in [59]–[68] for some model nonlinear autonomous wave equations (see the survey [60]): then the asymptotics of type (7.19) hold in appropriate *local energy seminorms*. However, the proof of the transitions (7.19) for the coupled equations (7.14) is still an open problem.

Comment 7.5 The nonlinear eigenvalue problem (7.16) could be handled by the iterative procedure. Namely, first let us neglect the fields $\phi_\omega(\mathbf{x}), \mathbf{A}_\omega(\mathbf{x})$ in the Schrödinger equation of (7.16) and get the first *Born approximation*

$$(7.20) \quad \omega\hbar\psi_\omega(\mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{x})]^2\psi_\omega(\mathbf{x}) + e\phi^{\text{ext}}(\mathbf{x})\psi_\omega(\mathbf{x})$$

which is the linear eigenvalue problem of type (6.5). Further we could express $\phi_\omega(\mathbf{x}), \mathbf{A}_\omega(\mathbf{x})$ from the Poisson equations in (7.16) (neglecting $\mathbf{A}_\omega(\mathbf{x})$ in the RHS for the first approximation), substitute them into the first equation, and continue by induction. The "corrections" $\phi_\omega(\mathbf{x}), \mathbf{A}_\omega(\mathbf{x})$ are small in a sense since the RHS of the Poisson equations contain the small factors e and e/c . Hence we could expect that the iterations converge, and even the solution of the linear problem (7.20) give a good (first) approximation to a solution of the nonlinear problem (7.16). If the iterations converge, the limit functions obviously give one of the solutions to (7.16).

Comment 7.6 i) The (orthodox) *Copenhagen interpretation* of Quantum Mechanics (see [74]) postulates that each observation of a quantum observable $\hat{\mathbf{M}}$ gives one of its eigenvalue *with the probability one*.

ii) This postulate agrees with our previous comment in the case when $\hat{\mathbf{M}}$ commutes with the Schrödinger operator \mathcal{H} (with the external potentials $\phi^{\text{ext}}(\mathbf{x})$ and $\mathbf{A}^{\text{ext}}(\mathbf{x})$): in particular, this is true for the energy $\hat{\mathcal{E}} = \mathcal{H}$. Namely, between two successive macroscopic observations, the quantum system could be considered as unperturbed even if the observation provide a disturbance. Therefore, the attraction (7.19) implies that before the observation the wave function $\psi(t) := \psi(\cdot, t)$ is "almost sure" close to a solitary wave $\psi_\omega e^{-i\omega t}$. Furthermore, $\psi_\omega \approx \psi_k$, where ψ_k is an eigenfunction of the linear operator \mathcal{H} (see previous comment). On the other hand, ψ_k is also an eigenfunction of the observable $\hat{\mathbf{M}}$. Hence the (mean) value is equal to $\mathbf{M}(t) := \langle \psi(t), \hat{\mathbf{M}}\psi(t) \rangle \approx \langle \psi_k(t), \hat{\mathbf{M}}\psi_k(t) \rangle$ that coincides with the corresponding eigenvalue.

iii) In the case when $\hat{\mathbf{M}}$ does not commute with the Schrödinger operator \mathcal{H} , this interpretation fails. One could expect then that the measuring process should modify the Schrödinger operator of the combined system (observable system + measuring instrument) to insure the commutation.

7.4 Charge Continuity Equations

The Lagrangian density (7.3) is invariant with respect to the internal rotations (39.12). Furthermore, the densities ρ, \mathbf{j} from (7.4) coincide with the corresponding Noether currents of the form (40.23) up to a factor. Therefore, the Noether theorem II implies

Lemma 7.7 *For any solution to the Schrödinger equation*

$$(7.21) \quad [i\hbar\partial_t - e\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi(t, \mathbf{x}),$$

the corresponding charge-current densities ρ, \mathbf{j} satisfy the continuity equation (4.2).

Let us prove a more general relation:

Lemma 7.8 *For any two solutions $\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x})$ to the Schrödinger equation in (7.21) the following identity holds:*

$$(7.22) \quad \dot{\rho}_{12}(t, \mathbf{x}) + \text{div } \mathbf{j}_{12}(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

where

$$(7.23) \quad \left\{ \begin{array}{l} \rho_{12}(t, \mathbf{x}) := e\psi_1(t, \mathbf{x})\bar{\psi}_2(t, \mathbf{x}), \\ \mathbf{j}_{12}(t, \mathbf{x}) := \frac{e}{2\mu} \left([-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]\psi_1(t, \mathbf{x}) \right) \bar{\psi}_2(t, \mathbf{x}) \\ \quad + \frac{e}{2\mu} \left([i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]\bar{\psi}_2(t, \mathbf{x}) \right) \psi_1(t, \mathbf{x}). \end{array} \right.$$

Remark 7.9 For $\psi_1 = \psi_2$ the expressions (7.23) become (7.2) by (2.3), and the identity (7.22) becomes (4.2).

Proof of Lemma 7.8 Let us write the equation (7.21) for ψ_1 and $\bar{\psi}_2$:

$$(7.24) \quad \left\{ \begin{array}{l} [i\hbar\partial_t - e\phi(\mathbf{x})]\psi_1(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi_1(t, \mathbf{x}), \\ [-i\hbar\partial_t - e\phi(\mathbf{x})]\bar{\psi}_2(t, \mathbf{x}) = \frac{1}{2\mu}[i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\bar{\psi}_2(t, \mathbf{x}) \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Let us multiply the first equation by $i\bar{\psi}_2(t, \mathbf{x})$ and add the second equation multiplied by $-i\psi_1(t, \mathbf{x})$. Then we get,

$$(7.25) \quad \begin{aligned} & i\hbar\partial_t\psi_1(t, \mathbf{x})i\bar{\psi}_2(t, \mathbf{x}) + i\hbar\partial_t\bar{\psi}_2(t, \mathbf{x})i\psi_1(t, \mathbf{x}) \\ &= \frac{1}{2\mu} \left([-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\psi_1(t, \mathbf{x}) \right) i\bar{\psi}_2(t, \mathbf{x}) \\ & \quad - \frac{1}{2\mu} \left([i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\bar{\psi}_2(t, \mathbf{x}) \right) i\psi_1(t, \mathbf{x}). \end{aligned}$$

This may be rewritten as

$$(7.26) \quad \begin{aligned} & -\partial_t[\psi_1(t, \mathbf{x})\bar{\psi}_2(t, \mathbf{x})] \\ &= \frac{1}{2\mu} \nabla \cdot \left\{ \left([-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]\psi_1(t, \mathbf{x}) \right) \bar{\psi}_2(t, \mathbf{x}) \right. \\ & \quad \left. + \left([i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]\bar{\psi}_2(t, \mathbf{x}) \right) \psi_1(t, \mathbf{x}) \right\}. \quad \blacksquare \end{aligned}$$

Exercise 7.10 Check the identity (7.26).

7.5 Born Approximation

Let us assume that the *induced Maxwell field* $\phi(t, \mathbf{x})$, $\mathbf{A}(t, \mathbf{x})$ in (7.4) is small compared to the external field. Then we can neglect the induced field in the first equation and consider the approximate equation (cf. (7.20))

$$(7.27) \quad [i\hbar\partial_t - e\phi^{\text{ext}}(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(t, \mathbf{x})]^2\psi(t, \mathbf{x}).$$

Its solution is the *Born approximation* to the wave field ψ in (7.4). Let us solve the equation and substitute the solution ψ into the RHS of the second equation of (7.4) also neglecting the induced Maxwell field:

$$(7.28) \quad \frac{1}{4\pi} \nabla_{\alpha} \mathcal{F}^{\alpha\beta}(t, \mathbf{x}) = \left(\begin{array}{l} \rho := e|\psi(t, \mathbf{x})|^2 \\ \frac{\mathbf{j}\beta}{c} := \frac{e}{\mu c}[-i\hbar\nabla_{\beta} - \frac{e}{c}\mathbf{A}_{\beta}^{\text{ext}}(t, \mathbf{x})]\psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}) \end{array} \right).$$

Then the solution $\mathcal{F}^{\alpha\beta}(t, \mathbf{x})$ is the *Born approximation* to the Maxwell field radiated by the atom. This process of approximations can be iterated.

Part III

Application of Schrödinger Theory

8 Spectrum of Hydrogen Atom

For a hydrogen atom we study the corresponding Schrödinger equation and find the spectrum as the solution to the eigenvalue problem. The proof is based on splitting the space $L^2(S^2)$ in a sum of orthogonal eigenspaces of the spherical Laplacian.

Let us consider the hydrogen atom, which has precisely one electron of negative charge e . The Rutherford experiment shows that the positive charge $-e$ is concentrated in a very small region called “nucleus”, so its Maxwell field is Coulombic $\phi(t, \mathbf{x}) = -e/|\mathbf{x}|$. We assume the magnetic field of the nucleus to be zero: $\mathbf{A}(t, \mathbf{x}) = 0$. Then the Schrödinger equation (6.2) becomes

$$(8.1) \quad i\hbar\partial_t\psi(t, \mathbf{x}) = \mathcal{H}\psi(t, \mathbf{x}) := -\frac{\hbar^2}{2\mu}\Delta\psi(t, \mathbf{x}) - \frac{e^2}{|\mathbf{x}|}\psi(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

and the stationary Schrödinger equation (6.9) becomes,

$$(8.2) \quad E_\omega\psi_\omega(\mathbf{x}) = \mathcal{H}\psi(\mathbf{x})_\omega.$$

Theorem 8.1 *The quantum stationary states $\psi_\omega \in \mathcal{E} := L^2(\mathbb{R}^3)$ of the hydrogen atom exist for energies $E_\omega = E_n := -2\pi\hbar R/n^2$, where $R := \mu e^4/(4\pi\hbar^3)$ is the **Rydberg constant** and $n = 1, 2, 3, \dots$. For other energies, stationary states do not exist.*

We will prove the theorem in this lecture.

8.1 Spherical Symmetry and Separation of Variables

Rotation invariance

A basic issue in solving the eigenvalue problem (8.2) is its spherical symmetry which implies the angular momentum conservation. Namely, the Schrödinger operator \mathcal{H} is invariant with respect to all rotations of the space \mathbb{R}^3 since the Laplacian Δ and the Coulomb potential are invariant under the rotations. The Schrödinger operator \mathcal{H} is invariant with respect to all rotations of the space \mathbb{R}^3 . This means that for any $k = 1, 2, 3$,

$$(8.3) \quad \mathcal{H}\hat{R}_k(\varphi) = \hat{R}_k(\varphi)\mathcal{H}, \quad \varphi \in \mathbb{R},$$

where $(\hat{R}_k(\varphi)\psi)(\mathbf{x}) := \psi(R_k(\varphi)\mathbf{x})$ and $R_k(\varphi)$ is a space rotation around the unit vector \mathbf{e}_k with an angle of φ radian in a positive direction, $\mathbf{e}_1 = (1, 0, 0)$, etc. The commutation holds by (8.2) since the Laplacian Δ and the Coulomb potential are invariant under all rotations of the space \mathbb{R}^3 . Differentiating (8.3) in φ at $\varphi = 0$, we get

$$(8.4) \quad [\mathcal{H}, \nabla_{\varphi_k}] = 0, \quad k = 1, 2, 3,$$

where

$$\nabla_{\varphi_k} := \left. \frac{d}{d\varphi} \right|_{\varphi=0} \hat{R}_k(\varphi).$$

Then \mathcal{H} also commutes with $\mathbf{H}_k := -i\nabla_{\varphi_k}$ and the angular momentum operators $\hat{\mathbf{L}}_k$ since $\hat{\mathbf{L}}_k = \hbar\mathbf{H}_k$ by (6.19):

$$(8.5) \quad [\mathcal{H}, \mathbf{H}_k] = 0, \quad [\mathcal{H}, \hat{\mathbf{L}}_k] = 0, \quad k = 1, 2, 3, .$$

Hence the angular momenta $\mathbf{L}_k(t) := \langle \psi(t), \hat{\mathbf{L}}_k \psi(t) \rangle$ are conserved.

Remark 8.2 *Let us note that the angular momentum conservation played a crucial role in the determination of the hydrogen spectrum in the Bohr-Sommerfeld “old quantum theory” (see Exercise 5).*

Spherical harmonics

Now we can explain our general strategy in proving Theorem 8.1. Namely, the commutation (8.4) provides a solution of the spectral problem (8.2) by a separation of variables. More precisely, the strategy relies on the following three general arguments:

I. The commutation (8.5) obviously implies that the operator $\mathbf{H}^2 := \mathbf{H}_1^2 + \mathbf{H}_2^2 + \mathbf{H}_3^2$ commutes with \mathcal{H} by (8.5):

$$(8.6) \quad [\mathcal{H}, \mathbf{H}^2] = 0.$$

Hence, each eigenspace of the Schrödinger operator \mathcal{H} is invariant with respect to each operators \mathbf{H}_k and \mathbf{H}^2 . Moreover, the operator \mathbf{H}^2 commutes also with each operator \mathbf{H}_k , for example,

$$(8.7) \quad [\mathbf{H}_k, \mathbf{H}^2] = 0, \quad k = 1, 2, 3.$$

Exercise 8.3 Check (8.7). **Hint:** First, prove the commutation relations $[\mathbf{H}_k, \mathbf{H}_j] = -i\hbar\epsilon_{kjl}\mathbf{H}_l$ where ϵ_{kjl} is a totally antisymmetric tensor.

Since the operators \mathbf{H}_k , $k = 1, 2, 3$, do not commute, they cannot be diagonalized simultaneously. On the other hand, the operators \mathcal{H} , \mathbf{H}^2 and, for example, \mathbf{H}_3 , commute with each other. Hence, we could expect that there is a basis of common eigenfunctions for them.

II. First, let us diagonalize simultaneously \mathbf{H}_3 and \mathbf{H}^2 . We consider the eigenvalue problem (8.2) in the Hilbert space $\mathcal{E} := L^2(\mathbb{R}^3)$. On the other hand, both operators, \mathbf{H}_3 and \mathbf{H}^2 , act only on the angular variables in spherical coordinates. Hence, the operators act also in the Hilbert space $\mathcal{E}_1 := L^2(S, dS)$, where S stands for the two-dimensional sphere $|\mathbf{x}| = 1$.

Theorem 8.4 *i) In \mathcal{E}_1 there exist an orthonormal basis of Spherical Harmonics $Y_l^m(\theta, \varphi)$ which are common eigenfunctions of \mathbf{H}_3 and \mathbf{H}^2 :*

$$(8.8) \quad \mathbf{H}_3 Y_l^m = m Y_l^m, \quad \mathbf{H}^2 Y_l^m = l(l+1) Y_l^m, \quad m = -l, -l+1, \dots, l,$$

where $l = 0, 1, 2, \dots$

ii) $Y_l^m(\theta, \varphi) = F_l^m(\theta)e^{im\varphi}$, where $F_l^m(\theta)$ are real functions.

We will prove this theorem in Section 10. It is important that each space of the common eigenfunctions is one-dimensional since the eigenvalues depend on l and m . This suggests that we could construct the eigenfunctions of the Schrödinger operator \mathcal{H} , by separation of variables, in the form

$$(8.9) \quad \psi(\mathbf{x}) = R(r)Y_l^m(\theta, \varphi).$$

The following theorem justify this choice of a particular form for the eigenfunctions.

Lemma 8.5 *(On Separation of Variables) Each solution to the spectral problem (8.2) is a sum (or a series) of the solutions of the particular form (8.9).*

Formal Proof Let π_l^m denote the orthogonal projection in \mathcal{E}_1 onto the linear span \mathcal{E}_l^m of Y_l^m . Let us define its action also in \mathcal{E} , by the formula

$$(8.10) \quad (\Pi_l^m \psi)(r, \cdot, \cdot) := \pi_l^m[\psi(r, \cdot, \cdot)], \quad r > 0,$$

in the spherical coordinates r, θ, φ . Then Π_l^m commutes with the Schrödinger operator \mathcal{H} since the latter commutes with \mathbf{H}_3 and \mathbf{H}^2 :

$$(8.11) \quad [\mathcal{H}, \Pi_l^m] = 0.$$

Hence, applying Π_l^m to (8.2), we get *formally*

$$(8.12) \quad E_\omega \Pi_l^m \psi_\omega = \mathcal{H} \Pi_l^m \psi_\omega.$$

It remains to note that

- i) The function $\Pi_l^m \psi_\omega$ has the form (8.9) since $\pi_l^m[\psi_\omega(r, \cdot, \cdot)] \in \mathcal{E}_l^m$, and the space \mathcal{E}_l^m is one-dimensional;
- ii) $\psi_\omega = \sum_{l,m} \Pi_l^m \psi_\omega$. ■

Remarks 8.6 *i) A complete solution to the spectral problem (8.8) relies on an investigation of all commutation relations of the operators \mathbf{H}_k , $k = 1, 2, 3$, i.e. the Lie algebra generated by them.*
ii) It remains still to determine the radial function in (8.9). We will substitute (8.9) into the equation (8.8). This gives a radial eigenvalue problem which will be solved explicitly.

8.2 Spherical Laplacian

To determine the radial functions in (8.9), let us express the Laplacian operator Δ in spherical coordinates r, θ, φ : by definition,

$$(8.13) \quad x_3 = r \cos \theta, \quad x_1 = r \sin \theta \cos \varphi, \quad x_2 = r \sin \theta \sin \varphi.$$

The operator Δ is symmetric in the real Hilbert space $L^2(\mathbb{R}^3)$. Hence, it is defined uniquely by the quadratic form $(\Delta\psi, \psi)$, where $\psi \in D := \{\psi \in L^2(\mathbb{R}^3) : \psi^{(\alpha)} \in L^2(\mathbb{R}^3) \cap C(\mathbb{R}^3), |\alpha| \leq 2\}$. In spherical coordinates

$$(8.14) \quad (\Delta\psi, \psi) = -(\nabla\psi, \nabla\psi) = -\int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi |\nabla\psi(r, \theta, \varphi)|^2 r^2 \sin \theta.$$

Geometrically it is evident that

$$(8.15) \quad \nabla\psi(r, \theta, \varphi) = \mathbf{e}_r \nabla_r \psi + \mathbf{e}_\theta \frac{\nabla_\theta \psi}{r} + \mathbf{e}_\varphi \frac{\nabla_\varphi \psi}{r \sin \theta},$$

where $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi$ are the orthogonal unit vectors proportional to $\nabla_r, \nabla_\theta, \nabla_\varphi$, respectively. Therefore, (8.14) becomes

$$(8.16) \quad (\Delta\psi, \psi) = -\int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi \left(|\nabla_r \psi|^2 + \left| \frac{\nabla_\theta \psi}{r} \right|^2 + \left| \frac{\nabla_\varphi \psi}{r \sin \theta} \right|^2 \right) r^2 \sin \theta.$$

Integrating by parts, we get

$$(8.17) \quad \begin{aligned} (\Delta\psi, \psi) &= \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi \left(r^{-2} \nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin \theta \nabla_\theta \psi}{r^2 \sin \theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2 \theta} \right) \psi r^2 \sin \theta \\ &= \left(r^{-2} \nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin \theta \nabla_\theta \psi}{r^2 \sin \theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2 \theta}, \psi \right). \end{aligned}$$

Therefore, we get the Laplacian operator in spherical coordinates,

$$(8.18) \quad \Delta\psi = r^{-2} \nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin \theta \nabla_\theta \psi}{r^2 \sin \theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2 \theta} = r^{-2} \nabla_r r^2 \nabla_r \psi + r^{-2} \Lambda \psi,$$

where Λ is the differential operator on the sphere S with coordinates θ, φ :

$$(8.19) \quad \Lambda = \frac{\nabla_\theta \sin \theta \nabla_\theta}{\sin \theta} + \frac{\nabla_\varphi^2}{\sin^2 \theta},$$

Exercise 8.7 Check the integration by parts (8.17).

Definition 8.8 The operator Λ is called **Spherical Laplacian Operator**.

Exercise 8.9 Check the identity

$$(8.20) \quad \Lambda = -\mathbf{H}^2.$$

Hint: Both operators are second order spherically symmetric elliptic operators.

8.3 Radial Equation

Here we deduce Theorem 8.1 from Theorem 8.4. Let us consider a nonzero finite energy solution $\psi_\omega(\mathbf{x}) \in \mathcal{E} := L^2(\mathbb{R}^3)$ to the problem (8.2) of the form (8.9). Substituting into (8.8), we get by (8.18) and (8.20) that

$$(8.21) \quad -\frac{2\mu E_\omega}{\hbar^2} R(r) = r^{-2} \nabla_r r^2 \nabla_r R(r) - \frac{l(l+1)}{r^2} R(r) + \frac{2\mu e^2}{\hbar^2 r} R(r), \quad r > 0.$$

For $r \rightarrow \infty$, the equation becomes

$$(8.22) \quad -\frac{2\mu E_\omega}{\hbar^2} R(r) \sim R_l''(r).$$

This suggests that $E_\omega < 0$ and the asymptotics $R(r) \sim e^{-\gamma r}$, $r \rightarrow \infty$ holds, where

$$(8.23) \quad \gamma = \sqrt{-2\mu E_\omega}/\hbar > 0.$$

Next, let us write $R(r) = e^{-\gamma r} F(r)$. Substituting into (8.21), we get

$$(8.24) \quad F'' + \left[\frac{2}{r} - 2\gamma \right] F' + \left[\frac{d}{r} - \frac{l(l+1)}{r^2} \right] F = 0, \quad r > 0,$$

where $d = b - 2\gamma$ with $b = 2\mu e^2/\hbar^2$. Finally, let us introduce the new variable $\rho = 2\gamma r$, then (8.24) becomes

$$(8.25) \quad f'' + \left[\frac{2}{\rho} - 1 \right] f' + \left[\frac{\lambda - 1}{\rho} - \frac{l(l+1)}{\rho^2} \right] f = 0, \quad \rho > 0,$$

where $f(\rho) = F(r)$ and $\lambda = b/(2\gamma)$. Now let us seek for a solution f of the form

$$(8.26) \quad f(\rho) = \rho^s (a_0 + a_1 \rho + a_2 \rho^2 + \dots) \equiv L(\rho) \rho^s$$

with $a_0 \neq 0$. We will find two linearly independent solutions: one with $s \geq 0$ and another with $s \leq -1$. Only the solution with $s \geq 0$ is appropriate. For $s \leq -1$ the corresponding eigenfunction $\psi_\omega(x)$ is not a function of finite energy since $\nabla \psi_\omega(x) \notin L^2(\mathbb{R}^3)$ and $\int \phi(\mathbf{x}) |\psi(\mathbf{x})|^2 d\mathbf{x} = \infty$. Substituting (8.26) into (8.25), we get

$$(8.27) \quad \rho^2 L'' + \left[2s\rho + \left[\frac{2}{\rho} - 1 \right] \rho^2 \right] L' + \left[s(s-1) + \left[\frac{2}{\rho} - 1 \right] s\rho + \left[\frac{\lambda-1}{r} - \frac{l(l+1)}{\rho^2} \right] \rho^2 \right] L = 0$$

for $\rho > 0$. After some evaluation, we have

$$(8.28) \quad \rho^2 L'' + \rho[2(s+1) - \rho] L' + [\rho(\lambda - 1 - s) + s(s+1) - l(l+1)] L = 0, \quad \rho > 0.$$

Setting $\rho = 0$, we get *formally* that $s(s+1) - l(l+1) = 0$. Hence, $s = l$ since $-l-1 \leq -1$. With $s = l$, this equation becomes

$$(8.29) \quad \rho L'' + [2(l+1) - \rho] L' + [\lambda - 1 - l] L = 0, \quad \rho > 0.$$

Let us substitute here $L(\rho) = a_0 + a_1\rho + a_2\rho^2 + \dots$ and equate the coefficients with identical powers of ρ :

$$(8.30) \quad \begin{cases} \rho^0 : & 2(l+1) + (\lambda - 1 - l)a_0 = 0, \\ \rho^1 : & 2a_2 + 2(l+1)2a_2 - a_1 + (\lambda - 1 - l)a_1 = 0, \\ \rho^2 : & 3 \cdot 2a_3 + 2(l+1)3a_3 - 2a_2 + (\lambda - 1 - l)a_2 = 0, \\ \dots & \\ \rho^k : & (k+1)ka_{k+1} + 2(l+1)(k+1)a_{k+1} - ka_k + (\lambda - 1 - l)a_k = 0, \\ \dots & \end{cases}$$

Therefore, we get the recursive relation

$$(8.31) \quad a_{k+1} = \frac{k - (\lambda - 1 - l)}{(k+1)(k+2l+2)} a_k.$$

This implies $a_{k+1}/a_k \sim 1/k$ if all $a_k \neq 0$. Then we have $|L(\rho)| \geq Ce^\rho = Ce^{2\gamma r}$ with a $C > 0$. Hence, $R(r) = F(r)e^{-\gamma r} \rightarrow \infty$ as $r \rightarrow \infty$, which contradicts the definition of a quantum stationary state. Therefore, $a_{k+1} = 0$ and $a_k \neq 0$ for some $k = 0, 1, 2, \dots$. Then $k - (\lambda - 1 - l) = 0$, so

$$(8.32) \quad \lambda = \frac{b}{2\gamma} = k + l + 1 = n = 1, 2, \dots$$

Substituting here $b = 2\mu e^2/\hbar^2$ and $\gamma = \sqrt{-2\mu E}/\hbar$, we get finally,

$$(8.33) \quad E = E_n := -\frac{2\pi\hbar R}{n^2}, \quad n = 1, 2, \dots,$$

where $R := \mu e^4/(4\pi\hbar^3)$. ■

Corollary 8.10 *The eigenfunctions ψ_ω for $\omega = -2\pi R/n^2 = -\mu e^4/(2\hbar^3 n^2)$ have the following form in spherical coordinates:*

$$(8.34) \quad \psi_\omega = \psi_{lmn} := Ce^{-r/a_n} P_l^n(r) F_l^m(\theta) e^{im\varphi}.$$

Here

$$(8.35) \quad a_n := 1/\gamma = \hbar/\sqrt{-2\mu E_n} = \hbar^2 n/(\mu e^2),$$

and $P_l^n(r)$ is a polynomial function of degree $n - 1 \geq l$, $F_l^m(\theta)e^{im\varphi} \in D(l)$ and $m = -l, \dots, l$.

Exercise 8.11 *Calculate the multiplicity of the eigenvalue E_n .*

Hint: *It is equal to $\sum_{0 \leq l \leq n-1} (2l+1)$.*

Remark 8.12 For the stationary state (8.34)

i) The function $L(\rho) = \rho^l(a_0 + \dots + a_k\rho^k)$ is a polynomial of degree $k + l = n - 1$. Hence, $n - 1$ equals the number of zeros of the function $L(\rho)$, $\rho \geq 0$, and n is called the *principal quantum number* of the stationary state.

ii) The component \mathbf{L}_3 of angular momentum is equal to $m\hbar$. This follows from (6.13), (6.17) and (6.19). The number m is called the *magnetic quantum number* of the stationary state since it is responsible for the change of the energy in the magnetic field (see (16.10)).

iii) The (squared) *absolute value of angular momentum* $\mathbf{L}^2 := \mathbf{L}_1^2 + \mathbf{L}_2^2 + \mathbf{L}_3^2$ is equal to $l(l+1)\hbar^2$. This follows from (8.8) since $\mathbf{L}^2 = \hbar^2 \mathbf{H}^2$. The number l is called the *azimuthal quantum number* of the stationary state.

9 Spherical Spectral Problem

We split the space $L^2(S^2)$ in a sum of orthogonal eigenspaces of the spherical Laplacian. The proof uses rotation invariance of the Schrödinger operator and classification of all irreducible representations of commutation relations for a Lie algebra of angular momenta.

9.1 Hilbert-Schmidt Argument

We start the proof of Theorem 8.4 with a diagonalization of the operator \mathbf{H}^2 . This lemma and the *elliptic theory* [86] imply

Lemma 9.1 *The operator $\mathbf{H}^2 := \mathbf{H}_1^2 + \mathbf{H}_2^2 + \mathbf{H}_3^2$ in the space $\mathcal{E}_1 := L^2(S, dS)$ is selfadjoint and admits the spectral resolution*

$$(9.1) \quad \mathcal{E}_1 = \bigoplus_{n=0}^{\infty} L(n),$$

where $L(n) \subset C^\infty(S)$ are **finite-dimensional** orthogonal subspaces of \mathcal{E}_1 , and $\mathcal{H}^2|_{L(n)} = \lambda_n$ with $\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$.

Proof *Step i)* Each \mathbf{H}_k is a *symmetric* operator in \mathcal{E}_1 . Indeed, the rotations $\hat{R}_k(\varphi)$ form a unitary group in $\mathcal{E}_1 = L^2(S, dS)$, hence its generators ∇_{φ_k} are skew-adjoint. Therefore, the operator \mathcal{H}^2 is a symmetric nonnegative operator in \mathcal{E}_1 .

Step ii) The operator \mathcal{H}^2 is a nonnegative elliptic second order operator on the sphere S . This follows from the formula (8.19) by the identity (8.20). Hence, the operator $\mathbf{H}^2 + 1 : H^2(S) \rightarrow H^0(S)$ is invertible, where $H^k(S)$ stand for the Sobolev spaces on the sphere S . Respectively, the operator $(\mathbf{H}^2 + 1)^{-1} : H^0(S) \rightarrow H^2(S)$ is selfadjoint and compact in $H^0(S) = L^2(S, dS)$ by the Sobolev embedding theorem. Hence, the resolution (9.1) exists by the Hilbert-Schmidt theorem for the operator $(\mathbf{H}^2 + 1)^{-1}$. ■

We will prove that we can choose the resolution (9.1) with $\dim L(n) = 2l + 1$ and $\lambda_n = l(l + 1)$, where $l = l(n) = 0, 1, 2, \dots$

Lemma 9.2 *All spaces $L(n)$ are invariant with respect to rotations of the sphere.*

Proof This follows from the rotation invariance of \mathbf{H}^2 .

Corollary 9.3 *All spaces $L(n)$ are invariant with respect to the operators \mathbf{H}_k , $k = 1, 2, 3$.*

9.2 Lie Algebra of Angular Momenta

The linear span $(\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)$ is a Lie algebra since the following commutation relations hold:

$$(9.2) \quad [\mathbf{H}_1, \mathbf{H}_2] = i\mathbf{H}_3, \quad [\mathbf{H}_2, \mathbf{H}_3] = i\mathbf{H}_1, \quad [\mathbf{H}_3, \mathbf{H}_1] = i\mathbf{H}_2.$$

Exercise 9.4 *Check the commutation relations. Hint: $\mathbf{H}_k = -i(\mathbf{x} \times \nabla_{\mathbf{x}})_k$.*

Lemma 9.5 *For each $k = 1, 2, 3$,*

i) All spaces $L(n)$ are invariant with respect to \mathbf{H}_k .

ii) $[\mathbf{H}_k, \mathbf{H}^2] = 0$.

Proof *ad i)* The invariance holds by Corollary 9.3.

ad ii) The commutation holds by (8.7). ■

Definition 9.6 $\mathbf{H}_\pm = \mathbf{H}_1 \pm i\mathbf{H}_2$.

Lemma 9.7 All spaces $L(n)$ are invariant with respect to \mathbf{H}_\pm , and

$$(9.3) \quad [\mathbf{H}_3, \mathbf{H}_\pm] = \pm \mathbf{H}_\pm,$$

$$(9.4) \quad \mathbf{H}^2 = \mathbf{H}_+\mathbf{H}_- + \mathbf{H}_3(\mathbf{H}_3 - 1) = \mathbf{H}_-\mathbf{H}_+ + \mathbf{H}_3(\mathbf{H}_3 + 1),$$

$$(9.5) \quad \mathbf{H}^2 = \frac{1}{2}(\mathbf{H}_+\mathbf{H}_- + \mathbf{H}_-\mathbf{H}_+) + \mathbf{H}_3^2.$$

Proof The invariance follows from Lemma 9.5 *i*), (9.3) and (9.4) follow from (9.2), (9.5) follows from (9.4). \blacksquare

9.3 Irreducible Representations of Commutation Relations

Here we give a complete classification of all possible triples of Hermitian operators satisfying the commutation relations (9.2). It implies a part of Theorem 8.4 for integer or half odd-integer values of l .

Proposition 9.8 [75] Let E be a nonzero finite-dimensional complex linear space with a Hermitian scalar product, and

i) linear Hermitian operators \mathbf{H}^k , $k = 1, 2, 3$, in E , satisfying the commutation relations (9.2).

ii) \mathbf{H}^2 is a scalar: $\mathbf{H}^2 = \alpha \in \mathbb{C}$,

iii) The space E is irreducible, i.e. it does not contain any nontrivial subspace which is invariant under all \mathbf{H}_k .

Then there exists a **spin number** $J = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ and an orthonormal basis in E , $\{e_m : m = -J, -J + 1, \dots, J - 1, J\}$, such that $\alpha = J(J + 1)$, and the action of the operators is given by

$$(9.6) \quad \mathbf{H}_1 e_m = \frac{s_{Jm}^+}{2} e_{m+1} + \frac{s_{Jm}^-}{2} e_{m-1}, \quad \mathbf{H}_2 e_m = \frac{s_{Jm}^+}{2i} e_{m+1} - \frac{s_{Jm}^-}{2i} e_{m-1}, \quad \mathbf{H}_3 e_m = m e_m,$$

where

$$(9.7) \quad s_{Jm}^\pm = \sqrt{(J \mp m)(J \pm m + 1)}.$$

Proof Set $\mathbf{H}_\pm := \mathbf{H}_1 \pm i\mathbf{H}_2$, then all relations (9.3)-(9.5) hold by (9.2), as in Lemma 9.5. Since \mathbf{H}_3 is a Hermitian operator, there exists a basis in E , provided by its eigenvectors: for each vector e_m of this basis we have $\mathbf{H}_3 e_m = m e_m$ with a real eigenvalue $m \in \mathbb{R}$.

Lemma 9.9 Either $e_\pm := \mathbf{H}_\pm e_m = 0$, or e_\pm is an eigenvector of \mathbf{H}_3 with the eigenvalue $m \pm 1$.

Proof (9.3) implies that $\mathbf{H}_3 e_\pm = \mathbf{H}_3 \mathbf{H}_\pm e_m = \mathbf{H}_\pm \mathbf{H}_3 e_m \pm \mathbf{H}_\pm e_m = m e_\pm \pm e_\pm$. \blacksquare

Further, (9.5) implies that α is real and $\alpha \geq m^2$ since $\mathbf{H}_\pm^* = \mathbf{H}_\mp$. Therefore, $\alpha = J(J + 1)$ for a unique $J \geq 0$. Finally, the identities (9.4) imply that $\mathbf{H}_-\mathbf{H}_- = \alpha - \mathbf{H}_3(\mathbf{H}_3 - 1)$ and $\mathbf{H}_+\mathbf{H}_+ = \alpha - \mathbf{H}_3(\mathbf{H}_3 + 1)$. Hence,

$$(9.8) \quad \begin{aligned} 0 \leq \|\mathbf{H}_- e_m\|^2 &= \langle e_m, \mathbf{H}_-^* \mathbf{H}_- e_m \rangle = \langle e_m, (\alpha - \mathbf{H}_3(\mathbf{H}_3 - 1)) e_m \rangle \\ &= [J(J + 1) - m(m - 1)] \|e_m\|^2 = [(J + m)(J - m + 1)] \|e_m\|^2, \end{aligned}$$

$$(9.9) \quad \begin{aligned} 0 \leq \|\mathbf{H}_+ e_m\|^2 &= \langle e_m, \mathbf{H}_+^* \mathbf{H}_+ e_m \rangle = \langle e_m, (\alpha - \mathbf{H}_3(\mathbf{H}_3 + 1)) e_m \rangle \\ &= [J(J + 1) - m(m + 1)] \|e_m\|^2 = [(J - m)(J + m + 1)] \|e_m\|^2. \end{aligned}$$

Therefore, $\mathbf{H}_-e_m = 0$ implies that either $m = -J$ or $m = J + 1$. But the last value is impossible since $J \geq 0$ and $J(J + 1) = \alpha \geq m^2$. Consequently, $m = -J \leq 0$. Similarly, $\mathbf{H}_+e_m = 0$ implies that $m = J \geq 0$. Therefore, m has to run in integral steps from $-J$ to J . Hence, $2J + 1 = 1, 2, \dots$. Further, $\{e_m : m = -J, -J + 1, \dots, J - 1, J\}$ is a basis since E is irreducible.

Finally, for the normalized vectors e_m , (9.8) and (9.9) imply that

$$(9.10) \quad \mathbf{H}_+e_m = s_{Jm}^+ e_{m+1}, \quad \mathbf{H}_-e_m = s_{Jm}^- e_{m-1}.$$

Hence, the formulas (9.6) follow. ■

Corollary 9.10 *Let us set $Z_{Jm}^+ = s_{J,-J}^+ \dots s_{J,m-1}^+$. Then the vectors $Y_J^m = \mathbf{H}_+^{J+m} e_{-J} / Z_{Jm}^+$, $m = -J, \dots, J$, constitute an orthonormal basis in the space E .*

Definition 9.11 *For $J = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, denote by $D(J)$ the space E endowed with the operators \mathbf{H}_k , $k = 1, 2, 3$, described by Proposition 9.8.*

Example 9.12 *For $J = \frac{1}{2}$, the operators \mathbf{H}_k in the orthonormal basis $(Y_{1/2}^{-1/2}, Y_{1/2}^{1/2})$ are represented by the matrices $\hat{\mathbf{s}}_k := \frac{1}{2}\sigma_k$, where σ_k are the Pauli matrices:*

$$(9.11) \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Exercise 9.13 *Deduce (9.11) from (9.6).*

9.4 Spherical Harmonics

Proposition 9.8 implies that each eigenspace $L(n)$ is isomorphic to a direct sum of a number $M(n, J)$ of spaces $D(J)$ with a unique value of $J = J(n) = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$. We can assume that λ_n are distinct for distinct n , then distinct values of J for a fixed n are impossible since the eigenvalue of \mathbf{H}^2 in $D(J)$, equal to $J(J + 1)$, is a strictly increasing function of $J \geq 0$. Hence, Theorem 8.4 i) will follow from the next lemma.

Lemma 9.14 *i) For every half-integer J we have $M(n, J) = 0$ for each n .*

ii) For every integer $l = 1, 2, 3, \dots$, there exists a unique $n = n(l)$ such that $M(n, l) = 1$.

Proof *i)* Each eigenspace $L(n)$ is a direct sum of the irreducible subspaces $D(J)$. Let us consider one of the subspaces. For each eigenvector $e_m \in D(J)$, of the operator \mathbf{H}_3 , we have the following differential equations,

$$(9.12) \quad \mathbf{H}_3 e_m(\theta, \varphi) = -i \nabla_\varphi e_m(\theta, \varphi) = m e_m(\theta, \varphi).$$

Therefore, $e_m(\theta, \varphi) = F_l^m(\theta) e^{im\varphi}$, so m is an integer number since the function $e_m(\theta, \varphi)$ has to be single-valued. Hence all possible J also are only integer: $J = l$.

ii) Let us consider the lowest eigenvalue $m = -l$ in the space $D(l)$. Then $\mathbf{H}_-e_{-l}(\theta, \varphi) = 0$ by Lemma 9.9. We will prove below that in spherical coordinates,

$$(9.13) \quad \mathbf{H}_- = -e^{-i\varphi} [\nabla_\theta - i \cot \theta \nabla_\varphi], \quad \mathbf{H}_+ = e^{i\varphi} [\nabla_\theta + i \cot \theta \nabla_\varphi].$$

Taking into account that $e_{-l}(\theta, \varphi) = F_l^{-l}(\theta) e^{-il\varphi}$, we get from $\mathbf{H}_-e_{-l}(\theta, \varphi) = 0$ the differential equation $(\nabla_\theta - l \cot \theta) F_l^{-l} = 0$. Hence, $F_l^{-l} = C \sin^l \theta$ which means that $-l$ is a simple eigenvalue. Hence, $M(n, l) \leq 1$ for all n .

It remains to check that $M(n, l) \neq 0$ for some n . This is equivalent to the existence of an eigenvector of the operator \mathbf{H}^2 with eigenvalue $l(l+1)$. However, this is obvious for the function $e_{-l}(\theta, \varphi)$. Namely, Eqn. (9.4) implies that

$$\mathbf{H}^2 e_{-l} = \mathbf{H}_+ \mathbf{H}_- e_{-l} + \mathbf{H}_3 (\mathbf{H}_3 - 1) e_{-l} = l(l+1) e_{-l}$$

since $\mathbf{H}_- e_{-l} = 0$ and $\mathbf{H}_3 e_{-l} = -l e_{-l}$. ■

Now Theorem 8.4 i) is proved. Further, Corollary 9.10 implies that the functions $\mathbf{H}_+^{l+m} e_{-l} / Z_{lm}^+$, $m = -l, \dots, l$, provide an orthonormal basis in the space $D(l)$. This implies Theorem 8.4 ii) since

$$(9.14) \quad \mathbf{H}_+^{l+m} e_{-l} = \left(e^{i\varphi} [\nabla_\theta + i \cot \theta \nabla_\varphi] \right)^{l+m} (\sin^l \theta e^{-il\varphi}) = F_l^m(\theta) e^{im\varphi},$$

where $F_l^m(\theta)$ is a real function. ■

Definition 9.15 *The normalized functions $Y_l^m(\theta, \varphi) = F_l^m(\theta) e^{im\varphi} / Z_{lm}^+$ are called **spherical harmonics**.*

9.5 Angular Momenta in Spherical Coordinates

We prove (9.13). First, let us rewrite (8.15) as

$$(9.15) \quad \nabla \psi(r, \theta, \varphi) = \mathbf{e}_r \nabla_r \psi + \mathbf{e}_\theta \frac{\nabla_\theta \psi}{r} + \mathbf{e}_\varphi \frac{\nabla_\varphi \psi}{r \sin \theta} = \mathbf{e}_1 \nabla_1 \psi + \mathbf{e}_3 \nabla_3 \psi + \mathbf{e}_3 \nabla_3 \psi,$$

where $\mathbf{e}_1 := (1, 0, 0)$, etc. Then it is evident geometrically that

$$(9.16) \quad \begin{cases} \mathbf{e}_r = (\mathbf{e}_1 \cos \varphi + \mathbf{e}_2 \sin \varphi) \sin \theta + \mathbf{e}_3 \cos \theta, \\ \mathbf{e}_\theta = (\mathbf{e}_1 \cos \varphi + \mathbf{e}_2 \sin \varphi) \cos \theta - \mathbf{e}_3 \sin \theta, \\ \mathbf{e}_\varphi = \mathbf{e}_2 \cos \varphi - \mathbf{e}_1 \sin \varphi. \end{cases}$$

Substituting this into (9.15), we get

$$(9.17) \quad \begin{cases} \nabla_1 = \sin \theta \cos \varphi \nabla_r + \cos \theta \cos \varphi \frac{\nabla_\theta}{r} - \sin \varphi \frac{\nabla_\varphi}{r \sin \theta}, \\ \nabla_2 = \sin \theta \sin \varphi \nabla_r + \cos \theta \sin \varphi \frac{\nabla_\theta}{r} + \cos \varphi \frac{\nabla_\varphi}{r \sin \theta}, \\ \nabla_3 = \cos \theta \nabla_r - \sin \varphi \frac{\nabla_\theta}{r}. \end{cases}$$

Substituting (9.17) and (8.13) into $\mathbf{H}_k = -i(\mathbf{x} \times \nabla_{\mathbf{x}})_k$, we get

$$(9.18) \quad \begin{cases} \mathbf{H}_1 = i(\sin \varphi \nabla_\theta + \cot \theta \cos \varphi \nabla_\varphi) \\ \mathbf{H}_2 = i(-\cos \varphi \nabla_\theta + \cot \theta \sin \varphi \nabla_\varphi) \\ \mathbf{H}_3 = -i \nabla_\varphi. \end{cases}$$

At last, the first two formulas imply (9.13). ■

10 Atom Dipole Radiation and Selection Rules

We evaluate a radiation of an atom with an electrostatic potential $\phi^{\text{ext}}(t, \mathbf{x}) = \phi^{\text{ext}}(\mathbf{x})$ in a static external magnetic field with vector potential $\mathbf{A}^{\text{ext}}(\mathbf{x})$. Then the Schrodinger equation (7.27) becomes

$$(10.1) \quad [i\hbar\partial_t - e\phi^{\text{ext}}(\mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{x})]^2\psi(t, \mathbf{x}).$$

We apply the Bohr approximation and derive the Rydberg-Ritz combination principle and selection rules in the case of cylindric symmetry.

10.1 Rydberg-Ritz Combination Principle

Consider the atom in a statistical equilibrium with an appropriate environment. Then the wave function $\psi(t, \mathbf{x})$ admits an eigenfunction expansion

$$(10.2) \quad \psi(t, \mathbf{x}) = \sum_k c_k \psi_k(\mathbf{x}) e^{-i\omega_k t} + \int_0^\infty c_\omega \psi_\omega(\mathbf{x}) e^{-i\omega t} d\omega,$$

where ψ_k and ψ_ω are the eigenmodes of the time-independent Schrödinger equation belonging to the discrete and continuous spectrum, respectively. Further, the c_k and c_ω describe the degree of excitation of the individual eigenmodes.

Comment 10.1 The amplitudes c_k and c_ω are random variables corresponding to an equilibrium distribution. The equilibrium is maintained by an external source, like the electric discharge or heat bath, X-rays, etc. The equilibrium distribution depends on the accelerating voltage (or the heat bath temperature, X-rays wave-length, etc): the number of the excited modes generally increases with the voltage (or with the heat bath temperature, etc). A mathematical justification of the equilibrium distribution for the Maxwell-Schrödinger equations is an open problem. The justification of the convergence to equilibrium distribution in a thermal bath is done in [22]-[27] for a list of linear hyperbolic equations, and in [54] for a nonlinear one. The role of the thermal bath plays the initial wave field. The convergence to the equilibrium distribution is caused by the fluctuations of the initial field. The fluctuations are provided by the mixing condition of Rosenblatt or Ibragimov-Linnik type.

Let us analyze the Maxwell field produced by the corresponding charge and current densities from (7.28). The potentials of the field satisfy the Maxwell equations (4.6), (4.9) and their large time asymptotics for bounded $|x|$ is given by the retarded potentials (43.4):

$$(10.3) \quad \left\{ \begin{array}{l} \phi(t, \mathbf{x}) \sim \phi_{ret}(t, \mathbf{x}) := \int \frac{\rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \\ \mathbf{A}(t, \mathbf{x}) \sim \mathbf{A}_{ret}(t, \mathbf{x}) := \frac{1}{c} \int \frac{\mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \end{array} \right. \quad t \rightarrow \infty, \quad |\mathbf{x}| \leq R.$$

Let us assume that the atom is located at the point $\mathbf{y} = 0$. Then the densities $\rho(t, \mathbf{y})$, $\mathbf{j}(t, \mathbf{y})$ are localized in a ball $|\mathbf{y}| \leq a \ll 1$, where $a > 0$ is an *atom radius*, $a \sim 10^{-8}$ cm:

$$(10.4) \quad \rho(t, \mathbf{y}) = 0, \quad \mathbf{j}(t, \mathbf{y}) = 0, \quad |\mathbf{y}| > a.$$

Therefore, any macroscopic observation at a distance $|\mathbf{x}| \sim 1$ coincides with high precision with the *dipole approximation* (cf. (42.5))

$$(10.5) \quad \left\{ \begin{array}{l} \phi(t, \mathbf{x}) \sim \frac{1}{|\mathbf{x}|} \int_{|\mathbf{y}| \leq a} \rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y} \\ \mathbf{A}(t, \mathbf{x}) \sim \frac{1}{c|\mathbf{x}|} \int_{|\mathbf{y}| \leq a} \mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y} \end{array} \right. \quad t \rightarrow \infty, \quad |\mathbf{x}| \leq R.$$

Remark 10.2 *The approximation $\sim |\mathbf{x}|^{-1}$ corresponds to the Hertzian dipole radiation (45.67).*

Next, we calculate the charge-current densities corresponding to the wave function (10.2) in the Born approximation (7.28). We can write it using (2.3), in the form

$$(10.6) \quad \left\{ \begin{array}{l} \rho(t, \mathbf{y}) = e\psi(t, \mathbf{y})\overline{\psi(t, \mathbf{y})} \\ \mathbf{j}(t, \mathbf{y}) = \text{Re} \left[\left(\frac{e}{\mu} [-i\hbar \nabla_{\mathbf{y}} - \frac{e}{c} \mathbf{A}^{\text{ext}}(\mathbf{y})] \psi(t, \mathbf{y}) \right) \overline{\psi(t, \mathbf{y})} \right] \end{array} \right.$$

The integral over the continuous spectrum in the RHS of (10.2), is responsible only for the continuous spectrum of the atom radiation. Hence, the discrete spectrum is completely determined by the first sum. Substituting the sum into (10.6), we get the ‘discrete components’ of the charge and current densities in the form

$$(10.7) \quad \left\{ \begin{array}{l} \rho_d(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) \sim e \text{Re} \sum_{kk'} c_k \bar{c}_{k'} e^{-i(\omega_k - \omega_{k'})(t - |\mathbf{x} - \mathbf{y}|/c)} \psi_k(\mathbf{y}) \overline{\psi_{k'}(\mathbf{y})} \\ \mathbf{j}_d(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) \sim \\ \frac{e}{\mu} \text{Re} \sum_{kk'} c_k \bar{c}_{k'} e^{-i(\omega_k - \omega_{k'})t} \left([-i\hbar \nabla_{\mathbf{y}} - \frac{e}{c} \mathbf{A}^{\text{ext}}(\mathbf{y})] \psi_k(\mathbf{y}) e^{i\omega_k |\mathbf{x} - \mathbf{y}|/c} \right) \overline{\psi_{k'}(\mathbf{y})} e^{-i\omega_{k'} |\mathbf{x} - \mathbf{y}|/c}. \end{array} \right.$$

If we substitute this into (10.5), we get the following approximation for the discrete component of the radiation:

$$(10.8) \quad \left\{ \begin{array}{l} \phi(t, \mathbf{x}) \sim \frac{1}{|\mathbf{x}|} \text{Re} \sum_{kk'} \phi_{kk'}(\mathbf{x}) e^{-i(\omega_k - \omega_{k'})t} \\ \mathbf{A}(t, \mathbf{x}) \sim \frac{1}{c|\mathbf{x}|} \text{Re} \sum_{kk'} \mathbf{A}_{kk'}(\mathbf{x}) e^{-i(\omega_k - \omega_{k'})t} \end{array} \right. \quad t \rightarrow \infty, \quad |\mathbf{x}| \leq R.$$

Corollary 10.3 *The Rydberg-Ritz combination principle holds: the **discrete spectrum** of the atom radiation is contained in the set $\{\omega_{kk'} := \omega_k - \omega_{k'}\}$.*

Further, applying the formula $\text{Re} \sum_{kk'} S_{kk'} = \frac{1}{2} \sum_{kk'} (S_{kk'} + \bar{S}_{k'k})$, we get the following expressions for the limiting amplitudes in (10.8):

$$(10.9) \quad \left\{ \begin{array}{l} \phi_{kk'}(\mathbf{x}) \sim c_k \bar{c}_{k'} e \int_{|\mathbf{y}| \leq a} e^{i\omega_{kk'} |\mathbf{x} - \mathbf{y}|/c} \psi_k(\mathbf{y}) \overline{\psi_{k'}(\mathbf{y})} d\mathbf{y} \\ \mathbf{A}_{kk'}(\mathbf{x}) \approx c_k \bar{c}_{k'} \frac{e}{2\mu} \int_{|\mathbf{y}| \leq a} e^{i\omega_{kk'} |\mathbf{x} - \mathbf{y}|/c} \left\{ \left([-i\hbar \nabla_{\mathbf{y}} - \frac{e}{c} \mathbf{A}^{\text{ext}}(\mathbf{y})] \psi_k(\mathbf{y}) \right) \overline{\psi_{k'}(\mathbf{y})} \right. \\ \left. + \left([i\hbar \nabla_{\mathbf{y}} - \frac{e}{c} \mathbf{A}^{\text{ext}}(\mathbf{y})] \overline{\psi_{k'}(\mathbf{y})} \right) \psi_k(\mathbf{y}) \right\} d\mathbf{y} \end{array} \right.$$

with an error of order $\hbar\omega_k/c$ in the last formula. Further let us assume that

$$(10.10) \quad |\omega_{kk'}|a/c \ll 2\pi$$

which means that the wave length $\lambda_{kk'} := 2\pi c/|\omega_{kk'}| \gg a$.

Under condition (10.10) the exponent in the integrands of (10.9) is close to a constant $\exp(-i\omega_{kk'}|\mathbf{x}|/c)$ since $|\mathbf{y}| \leq a$. In addition, the ball of integration $|\mathbf{y}| \leq a$ can be substituted by the whole space since the eigenfunctions $\psi_k, \psi_{k'}$ are well localized. Therefore, (10.9) becomes

$$(10.11) \quad \left\{ \begin{array}{l} \phi_{kk'}(\mathbf{x}) \approx ec_k\bar{c}_{k'}e^{i\omega_{kk'}|\mathbf{x}|/c} \int \psi_k(\mathbf{y})\bar{\psi}_{k'}(\mathbf{y})d\mathbf{y} \\ \mathbf{A}_{kk'}(\mathbf{x}) \approx \frac{e}{2\mu}c_k\bar{c}_{k'}e^{i\omega_{kk'}|\mathbf{x}|/c} \int \left\{ \left([-i\hbar\nabla_{\mathbf{y}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{y})]\psi_k(\mathbf{y}) \right) \bar{\psi}_{k'}(\mathbf{y}) \right. \\ \left. + \left([i\hbar\nabla_{\mathbf{y}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{y})]\bar{\psi}_{k'}(\mathbf{y}) \right) \psi_k(\mathbf{y}) \right\} d\mathbf{y} \end{array} \right.$$

First formula of (10.11) implies that $\phi_{kk'}(\mathbf{x}) \approx 0$ for $\omega_{kk'} \neq 0$ by the orthogonality of different eigenfunctions. Hence, the radiation is represented by the magnetic potential

$$(10.12) \quad \mathbf{A}(t, \mathbf{x}) \sim \frac{1}{c|\mathbf{x}|} \frac{e}{2\mu} \text{Re} \sum_{kk'} c_k\bar{c}_{k'} e^{-i\omega_{kk'}(t-|\mathbf{x}|/c)} \mathbf{J}_{kk'}, \quad t \rightarrow \infty,$$

where $\mathbf{J}_{kk'}$ stands for the last integral in (10.11).

Note that the radiation field (10.12) is identical with the sum of the *Hertzian dipole radiation fields* of type (45.67) corresponding to the *dipole moments* $\sim \mathbf{J}_{kk'}$.

Comment 10.4 *The radiation field (10.12) brings the energy to infinity like the Hertzian dipole radiation field. Hence, it cannot be stationary without an external source of the energy (cf. Comment 10.1).*

The formulas (10.12) define the **intensity of the spectral line** $\omega_{kk'}$ corresponding to simple eigenvalues ω_k and $\omega_{k'}$:

$$(10.13) \quad \langle \mathbf{A}_{kk'} \rangle \sim \langle c_k\bar{c}_{k'} \rangle \mathbf{J}_{kk'},$$

where $\langle \cdot \rangle$ stands for the *mathematical expectation*. For the multiple eigenvalues the intensity is also proportional to the **multiplicity of the spectral line** which is defined by its splitting in a weak magnetic field (see [89]).

Traditional identification of the relative intensities (see [89]) means that

$$(10.14) \quad \frac{\langle \mathbf{A}_{kk'} \rangle}{\langle \mathbf{A}_{nn'} \rangle} = \frac{\langle c_k\bar{c}_{k'} \rangle \mathbf{J}_{kk'}}{\langle c_n\bar{c}_{n'} \rangle \mathbf{J}_{nn'}} \approx \frac{\mathbf{J}_{kk'}}{\mathbf{J}_{nn'}}.$$

It holds if $\langle c_k\bar{c}_{k'} \rangle$ varies slowly in $k, k' \in (N_1, N_2)$ with $N_2 - N_1$ sufficiently small.

10.2 Selection Rules for Cylindrical Symmetry

We consider the case of a **radial** electrostatic potential $\phi^{\text{ext}}(\mathbf{x}) = \phi^{\text{ext}}(|\mathbf{x}|)$ and a static **uniform** magnetic field B . Then the vector potential $\mathbf{A}^{\text{ext}}(\mathbf{x}) = B \times \mathbf{x}/2$. Let us consider a fixed spectral line $\omega_{kk'}$ corresponding to the frequencies ω_k and $\omega_{k'}$ and eigenfunctions (see (16.12))

$$(10.15) \quad \psi_{\omega} = R_{nl}(r)F_l^m(\theta)e^{im\varphi}, \quad \psi_{\omega'} = R_{n'l'}(r)F_{l'm'}(\theta)e^{im'\varphi}.$$

The next theorem demonstrates that the spectrum of the atom radiation is a very small subset of the set of all differences $\{\omega_{kk'} := \omega_k - \omega_{k'}\}$ from the Rydberg-Ritz combination principle.

Theorem 10.5 Let the condition (10.10) hold, $\phi^{\text{ext}}(\mathbf{x}) = \phi^{\text{ext}}(|\mathbf{x}|)$ and $\mathbf{A}^{\text{ext}}(\mathbf{x}) = B \times \mathbf{x}/2$. Then for the dipole approximations of the limiting amplitudes we have $\mathbf{J}_{kk'} = 0$ if either $l' \neq l \pm 1$ or $m' \neq m, m \pm 1$.

Proof Let us use the identity (7.22) for the solutions $\psi_k(\mathbf{x})e^{-i\omega_k t}$, $\psi_{k'}(\mathbf{x})e^{-i\omega_{k'} t}$ to Equation (7.27). Eqs. (7.23) give the densities

$$(10.16) \quad \left\{ \begin{array}{l} \rho_{kk'}(t, \mathbf{x}) := e^{-i\omega_{kk'} t} e\psi_k(\mathbf{x})\overline{\psi_{k'}(\mathbf{x})}, \\ \mathbf{j}_{kk'}(t, \mathbf{x}) := e^{-i\omega_{kk'} t} \left\{ \left([-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{x})]\psi_k(\mathbf{x}) \right) \overline{\psi_{k'}(\mathbf{x})} \right. \\ \left. + \left([i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(\mathbf{x})]\overline{\psi_{k'}(\mathbf{x})} \right) \psi_k(\mathbf{x}) \right\} \end{array} \right.$$

Then (7.22) becomes

$$(10.17) \quad -i\omega_{kk'}\rho_{kk'}(t, \mathbf{x}) + \text{div } \mathbf{j}_{kk'}(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

We see that the integral in the second formula of (10.11) is proportional to the integral of $\mathbf{j}_{kk'}(t, \mathbf{y})$:

$$(10.18) \quad \mathbf{J}_{kk'} \approx e^{i\omega_{kk'}|\mathbf{x}|/c} e^{i\omega_{kk'} t} \int \mathbf{j}_{kk'}(t, \mathbf{y}) d\mathbf{y}.$$

It remains to prove

Lemma 10.6

$$(10.19) \quad \int \mathbf{j}_{kk'}(t, \mathbf{x}) d\mathbf{x} = 0, \quad t \in \mathbb{R},$$

if either $l' \neq l \pm 1$ or $m' \neq m, m \pm 1$.

Proof Let us check the case $m' \neq m, m \pm 1$. The case $l' \neq l \pm 1$ will be considered in next section. Let us multiply (10.17) by the coordinate \mathbf{x}^p , $p = 1, 2, 3$, and integrate over \mathbb{R}^3 . Then we get by partial integration,

$$(10.20) \quad -i\omega_{kk'} \int \mathbf{x}^p \rho_{kk'}(t, \mathbf{x}) d\mathbf{x} - \int \mathbf{j}_{kk'}^p(t, \mathbf{x}) d\mathbf{x} = 0,$$

where $\mathbf{j}_{kk'}^p$ is the p -th component of $\mathbf{j}_{kk'}$. Hence,

$$(10.21) \quad \int \mathbf{j}_{kk'}^p(t, \mathbf{x}) d\mathbf{x} \sim c(t) \int \mathbf{x}^p \psi_k(\mathbf{x}) \cdot \psi_{k'}(\mathbf{x}) d\mathbf{x}.$$

Let us rewrite the last integral in spherical coordinates. Then we get by (10.15),

$$(10.22) \quad \begin{aligned} & \int \mathbf{x}^p \psi_k(\mathbf{x}) \cdot \psi_{k'}(\mathbf{x}) d\mathbf{x} \\ &= \int_0^\infty R_{nl}(r) \overline{R_{n'l'}(r)} r^3 dr \int_S \frac{\mathbf{x}^p}{r} F_l^m(\theta) e^{im\varphi} \overline{F_{l'm'}(\theta) e^{im'\varphi}} dS. \end{aligned}$$

Obviously, the last integral is equal to zero if $m' \neq m, m \pm 1$. ■

Remark 10.7 The RHS of (10.21) equals to the matrix element of the dipole moment operator $\hat{\mathbf{x}}^p$. Hence, the intensities $\mathbf{J}_{kk'}$ are proportional to the dipole moment like classical Hertzian dipole radiation (45.67). This is an example of the Bohr correspondence principle.

Comment 10.8 The second Bohr postulate states that the radiation of the spectral line $\omega_{kk'}$ is provided by the transition $\psi_k \mapsto \psi_{k'}$ between the stationary states. With this identification, the formulas (10.18) and (10.21) mean that the RHS of (10.21) is proportional to the probability of the transition.

10.3 Selection Rules for Orbital Momentum

The proof of the selection rule $l' = l \pm 1$ is more involved. We follow the method of [12, Appendix XXI]. First, let us re-express the $2l + 1$ eigenvectors $F_l^m(\theta)e^{im\varphi}$ of the spherical Laplacian with eigenvalue $-l(l + 1)$ in terms of real eigenvectors, e.g., like

$$(10.23) \quad \begin{aligned} Z_{l,m} &= F_l^m(\theta) \cos m\varphi \quad , \quad m = 0, \dots, l \\ Z_{l,m} &= F_l^m(\theta) \sin(m - l)\varphi \quad , \quad m = l + 1, \dots, 2l, \end{aligned}$$

then let us prove the following

Lemma 10.9 *The $Z_{l,m}$ may be expressed as $r^{-l}U_{l,m}(\mathbf{x})$, where each $U_{l,m}$ is a homogeneous polynomial of degree l in the variables x, y, z such that*

$$(10.24) \quad \Delta U_{l,m} = 0$$

holds.

Proof of Lemma 10.9

- 1) $r^{-l}U_{l,m}(\mathbf{x})$ only depends on the angular coordinates θ, φ .
- 2) From

$$\Delta = \frac{1}{r^2} \partial_r^2 + \frac{2}{r} \partial_r + \frac{1}{r^2} \Lambda$$

it follows that

$$(10.25) \quad 0 = \Delta U_{l,m} = r^{l-2} (l(l + 1)Z_{l,m} + \Lambda Z_{l,m})$$

therefore $Z_{l,m} \equiv r^{-l}U_{l,m}$ are eigenvectors of Λ with eigenvalues $-l(l + 1)$.

- 3) There are exactly $2l + 1$ linearly independent $U_{l,m}$ with the property $\Delta U_{l,m} = 0$. This may be shown as follows: There are $\frac{1}{2}(l + 2)(l + 1)$ linearly independent homogeneous polynomials of degree l , e.g., the monomials

$$x^l, (x^{l-1}y, x^{l-1}z), (x^{l-2}y^2, x^{l-2}yz, x^{l-2}z^2), \dots, (xy^{l-1}, \dots, xz^{l-1}), (y^l, \dots, z^l)$$

(here, all monomials with the same power of x are grouped together). Further, the Laplacian Δ acting on a general homogeneous polynomial V_l of degree l produces a general homogeneous polynomial \bar{V}_{l-2} of degree $l - 2$, $\Delta V_l = \bar{V}_{l-2}$. The condition that \bar{V}_{l-2} vanishes introduces therefore $\frac{1}{2}l(l - 1)$ conditions on the $\frac{1}{2}(l + 2)(l + 1)$ coefficients of V_l , therefore there exist

$$(10.26) \quad \frac{1}{2}(l + 2)(l + 1) - \frac{1}{2}l(l - 1) = 2l + 1$$

linearly independent V_l such that $\Delta V_l = 0$. Together with Theorem 8.4 this proves the above lemma. ■

From Theorem 8.4 (orthogonality of the $D(l)$ subspaces) it follows that

$$\int \sin \theta d\theta d\varphi Z_{l,m} Z_{l',m'} = 0$$

unless $l = l'$. Therefore, all that remains to prove is that

$$(10.27) \quad \frac{\mathbf{x}^p}{r} Z_{l,m} = Y_{l+1} + Y_{l-1}$$

where Y_{l+1} and Y_{l-1} are some general eigenvectors of $D(l+1)$ and $D(l-1)$, respectively, and \mathbf{x}^p is one component of the coordinate vector $\mathbf{x} = (x, y, z)$. Eq. (10.27) may be re-expressed, e.g., for $\mathbf{x}^p = x$, like

$$(10.28) \quad xU_l = U_{l+1} + r^2U_{l-1}$$

where the U_k are homogeneous polynomials of degree k which obey $\Delta U_k = 0$. Eq. (10.28) may be proved with the help of the following

Lemma 10.10 *Every homogeneous polynomial F_n of degree n may be uniquely expressed like*

$$(10.29) \quad F_n = U_n + r^2U_{n-2} + r^4U_{n-4} + \dots$$

where the U_k are homogeneous polynomials of degree k which obey $\Delta U_k = 0$.

The proof of this lemma will be given below. Certainly, xU_l of Eq. (10.28) is a homogeneous polynomial of degree $l+1$, therefore it may be re-expressed like

$$(10.30) \quad xU_l = U_{l+1} + r^2U_{l-1} + r^4U_{l-3} + \dots$$

Next, we act with the Laplace operator on both sides of Eq. (10.30). For the l.h.s. we get

$$(10.31) \quad \Delta xU_l = 2\partial_x U_l$$

and for the r.h.s. we use

$$(10.32) \quad \begin{aligned} \Delta \left(r^{2h} U_{l+1-2h} \right) &= r^{2h} \Delta U_{l+1-2h} + 2 \left(\nabla r^{2h} \right) \cdot \left(\nabla U_{l+1-2h} \right) + U_{l+1-2h} \Delta r^{2h} \\ &= 0 + 4hr^{2h-2} \mathbf{x} \cdot \nabla U_{l+1-2h} + 2h(2h+1)r^{2h-2} U_{l+1-2h} \\ &= 2h(2l+3-2h)r^{2h-2} U_{l+1-2h} \end{aligned}$$

where we used Euler's theorem for homogeneous polynomials,

$$\mathbf{x} \cdot \nabla U_n = nU_n.$$

Therefore, we get altogether

$$(10.33) \quad 2\partial_x U_l = 2(2l+1)U_{l-1} + 4(2l-1)r^2U_{l-3} + \dots$$

Now the important observation is that the l.h.s. vanishes upon a further application of the Laplace operator, $\Delta \partial_x U_l = \partial_x \Delta U_l = 0$, therefore the r.h.s. must vanish, as well, when the Laplace operator is applied. But, according to (10.32), this is possible only if all the U_{l+1-2h} except for U_{l-1} vanish (i.e., $U_{l-3} = U_{l-5} = \dots = 0$). This proves Eq. (10.28) and, therefore, the original statement that the integral (10.22) vanishes unless $l' = l \pm 1$. ■

We still have to prove Lemma (10.10).

Proof of Lemma 10.10 We prove the lemma by induction. It is obviously true for $n = 0, 1$. We assume that it is true for all $k < n$. Then it is also true for ΔF_n , which is a homogeneous polynomial of degree $n-2$, i.e.,

$$(10.34) \quad \Delta F_n = \bar{U}_{n-2} + r^2\bar{U}_{n-4} + \dots + r^{2h}\bar{U}_{n-2h} + \dots$$

The general solution to this equation, which is subject to the additional condition that it is a homogeneous polynomial of degree n , is

$$(10.35) \quad F_n = G_n + \frac{1}{2(2n-1)} r^2 \bar{U}_{n-2} + \dots + \frac{1}{2h(2n-2h+1)} r^{2h} \bar{U}_{n-2h}$$

as may be shown easily with the help of Eq. (10.32). Here G_n is an arbitrary homogeneous polynomial of degree n subject to the condition $\Delta G_n = 0$. By choosing

$$U_n = G_n, \quad U_{n-2h} = 2h(2n - 2h + 1)\bar{U}_{n-2h}$$

this can be brought into the form stated in Lemma (10.10). ■

11 Classical Scattering of Light: Thomson formula

Electromagnetic waves may be scattered by a charged matter. Here, we describe a light as a plane wave satisfying the free Maxwell equation and a matter as composed of classical, point-like particles obeying the Lorentz equation. The inconsistencies inherent to the concept of point-like charged particles are circumvented by applying a certain approximation (neglecting self-interactions). We apply the Hertzian dipole radiation formula and derive the Thomson differential cross-section.

11.1 Incident Plane Wave

J.C.Maxwell identified light with electromagnetic waves which are solutions to the Maxwell equations in *free space* with $\rho(t, \mathbf{x}) = 0$ and $\mathbf{j}(t, \mathbf{x}) = 0$. A general solution to the Maxwell equations is expressed through the potentials by (4.3):

$$(11.1) \quad \mathbf{E}(t, \mathbf{x}) = -\nabla_{\mathbf{x}}\phi(t, \mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}(t, \mathbf{x}), \quad \mathbf{B}(t, \mathbf{x}) = \text{rot } \mathbf{A}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

For the case $\rho(t, \mathbf{x}) = 0$ and $\mathbf{j}(t, \mathbf{x}) = 0$ the potentials $\mathbf{A}_0(t, \mathbf{x})$ and $\phi_0(t, \mathbf{x})$ are solutions to the homogeneous wave equations (4.9) and (4.6):

$$(11.2) \quad \square \mathbf{A}_0(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

$$(11.3) \quad \square \phi_0(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Let us choose $\phi(t, \mathbf{x}) = 0$ for concreteness, and

$$(11.4) \quad \mathbf{A}_0(t, \mathbf{x}) = A \sin k(x_3 - ct)(1, 0, 0)$$

with a wave number $k > 0$ and the corresponding frequency $\omega = kc > 0$. Then (11.1) gives

$$(11.5) \quad \mathbf{E}_0(t, \mathbf{x}) = kA \cos k(x_3 - ct)(1, 0, 0), \quad \mathbf{B}_0(t, \mathbf{x}) = kA \cos k(x_3 - ct)(0, 1, 0).$$

The corresponding energy flux (i.e., Poining vector) is

$$(11.6) \quad \mathbf{S}_0(t, \mathbf{x}) = \frac{c}{4\pi} \mathbf{E}_0(t, \mathbf{x}) \times \mathbf{B}_0(t, \mathbf{x}) = \frac{cE_0^2}{4\pi} \cos^2 k(x_3 - ct)(0, 0, 1), \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

where $E_0 := kA$. Let us note that the energy flux is directed along $\mathbf{e}_3 = (0, 0, 1)$ and its *intensity* is

$$(11.7) \quad I_0 := \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_0^T \mathbf{S}_0(t, \mathbf{x}) dt \right| = \frac{cE_0^2}{8\pi}.$$

11.2 Scattering Problem

We consider the scattering of the plane wave (11.5) by a classical electron. The scattering is described by the coupled Maxwell-Lorentz equations

$$(11.8) \quad \left\{ \begin{array}{l} \text{div } \mathbf{E}(t, \mathbf{x}) = 4\pi e \delta(\mathbf{x} - \mathbf{x}(t)), \quad \text{rot } \mathbf{E}(t, \mathbf{x}) = -\frac{1}{c} \dot{\mathbf{B}}(t, \mathbf{x}), \\ \text{div } \mathbf{B}(t, \mathbf{x}) = 0, \quad \text{rot } \mathbf{B}(t, \mathbf{x}) = \frac{1}{c} \dot{\mathbf{E}}(t, \mathbf{x}) + \frac{4\pi}{c} e \dot{\mathbf{x}} \delta(\mathbf{x} - \mathbf{x}(t)), \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

$$(11.9) \quad \mu \ddot{\mathbf{x}}(t) = e[\mathbf{E}(t, \mathbf{x}(t)) + \frac{1}{c} \dot{\mathbf{x}}(t) \times \mathbf{B}(t, \mathbf{x}(t))], \quad t \in \mathbb{R}.$$

The *free electron* is governed by the Lorentz equation (11.9) with the initial conditions

$$(11.10) \quad \mathbf{x}(t) = 0, \dot{\mathbf{x}}(t) = 0, \quad t < 0.$$

The incident plane wave appears in the initial conditions for the fields,

$$(11.11) \quad \mathbf{E}(t, \mathbf{x}) = \Theta(-x_3 + ct)\mathbf{E}_0(t, \mathbf{x}) - e\frac{\mathbf{x}}{|\mathbf{x}|^2}, \quad \mathbf{B}(t, \mathbf{x}) = \Theta(-x_3 + ct)\mathbf{B}_0(t, \mathbf{x}), \quad t < 0,$$

where Θ is the Heaviside function and $-e\mathbf{x}/|\mathbf{x}|$ is the static Coulomb field generated by the electron at the position (11.10). Let us note that the incident wave is a solution to the homogeneous Maxwell equations.

11.3 Neglecting the Self-Interaction

Let us split the solution to (11.8) like

$$(11.12) \quad \left\{ \begin{array}{l} \mathbf{E}(t, \mathbf{x}) = \Theta(-x_3 + ct)\mathbf{E}_0(t, \mathbf{x}) + \mathbf{E}_r(t, \mathbf{x}), \\ \mathbf{B}(t, \mathbf{x}) = \Theta(-x_3 + ct)\mathbf{B}_0(t, \mathbf{x}) + \mathbf{B}_r(t, \mathbf{x}), \end{array} \right. \quad t \in \mathbb{R},$$

where $\mathbf{E}_r(t, \mathbf{x}), \mathbf{B}_r(t, \mathbf{x})$ stand for the *radiated fields*. The fields are defined by the splitting. Then the Maxwell equations (11.8) read

$$(11.13) \quad \left\{ \begin{array}{l} \operatorname{div} \mathbf{E}_r(t, \mathbf{x}) = 4\pi e\delta(\mathbf{x} - \mathbf{x}(t)), \quad \operatorname{rot} \mathbf{E}_r(t, \mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}_r(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}_r(t, \mathbf{x}) = 0, \quad \operatorname{rot} \mathbf{B}_r(t, \mathbf{x}) = \frac{1}{c}\dot{\mathbf{E}}_r(t, \mathbf{x}) + \frac{4\pi}{c}e\dot{\mathbf{x}}\delta(\mathbf{x} - \mathbf{x}(t)), \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

since the incident wave in (11.11) is a solution to the homogeneous Maxwell equations. The initial conditions (11.11) become

$$(11.14) \quad \mathbf{E}_r(t, \mathbf{x}) = -e\frac{\mathbf{x}}{|\mathbf{x}|^2}, \quad \mathbf{B}_r(t, \mathbf{x}) = 0, \quad t < 0.$$

The Lorentz equation (11.9) now reads

$$(11.15) \quad \begin{aligned} \mu\ddot{\mathbf{x}}(t) &= e[\Theta(-\mathbf{x}_3(t) + ct)\mathbf{E}_0(t, \mathbf{x}(t)) + \mathbf{E}_r(t, \mathbf{x}(t))] \\ &+ \frac{1}{c}\dot{\mathbf{x}}(t) \times (\Theta(-\mathbf{x}_3(t) + ct)\mathbf{B}_0(t, \mathbf{x}(t)) + \mathbf{B}_r(t, \mathbf{x}(t))), \quad t \in \mathbb{R}. \end{aligned}$$

Unfortunately, the problem (11.13), (11.15) is not well posed. Namely, the solutions $\mathbf{E}_r(t, \mathbf{x}), \mathbf{B}_r(t, \mathbf{x})$ to (11.13) are infinite at the points $(t, \mathbf{x}(t))$, which is obvious from (11.14). Therefore, the RHS of equation (11.15) does not make sense for the fields $\mathbf{E}_r(t, \mathbf{x}), \mathbf{B}_r(t, \mathbf{x})$, which are solutions to (11.13).

Remark 11.1 *To make the problem well posed it is necessary to replace the point-like electron by the extended electron suggested by M. Abraham [1]. For this model, the well-posedness is proved in [67].*

Here we use another traditional approach to make the problem well posed, which is similar to the Born approximation. Concretely, we omit the radiation fields at the RHS of (11.15):

$$(11.16) \quad \mu\ddot{\mathbf{x}}(t) = e[\Theta(-\mathbf{x}_3 + ct)\mathbf{E}_0(t, \mathbf{x}) + \frac{1}{c}\dot{\mathbf{x}}(t) \times \Theta(-\mathbf{x}_3 + ct)\mathbf{B}_0(t, \mathbf{x})], \quad t \in \mathbb{R}.$$

Then we substitute the solution $\mathbf{x}(t)$ into the RHS of the Maxwell equations (11.13) to calculate the radiated fields $\mathbf{E}_r(t, \mathbf{x}), \mathbf{B}_r(t, \mathbf{x})$.

The approximation (11.16) means that we neglect the self-interaction of the electron, which can be justified for the “non-relativistic electron”. This means that the electron velocities are small compared to the speed of light:

$$(11.17) \quad \beta := \max_{t \in \mathbb{R}} |\dot{\mathbf{x}}(t)|/c \ll 1.$$

Then we can neglect the contribution of the magnetic field to the RHS of (11.16). Finally, we consider the equation

$$(11.18) \quad \mu \ddot{\mathbf{x}}(t) = e \mathbf{E}_0(t, \mathbf{x}), \quad t > 0.$$

This equation and the initial conditions (11.10) define the trajectory $\mathbf{x}(t)$ uniquely:

$$(11.19) \quad \mathbf{x}(t) = \frac{eA}{\mu k c^2} (1 - \cos kct)(1, 0, 0).$$

The condition (11.17) for the solution is equivalent to

$$(11.20) \quad \beta = \frac{|e|A}{\mu c^2} \ll 1.$$

Remark 11.2 *This relation means that the amplitude of the oscillations $\frac{|e|A}{\mu k c^2}$ is small compared to the wave length $2\pi/k$ of the incident wave.*

11.4 Scattering in Dipole Approximation

Now we have to solve the Maxwell equations (11.13) to define the radiation fields $\mathbf{E}_r(t, \mathbf{x}), \mathbf{B}_r(t, \mathbf{x})$. Our goal is an analysis of the energy flux at infinity, i.e., of the Poynting vector $\mathbf{S}_r(t, \mathbf{x}) := (c/4\pi)\mathbf{E}_r(t, \mathbf{x}) \times \mathbf{B}_r(t, \mathbf{x})$ as $|x| \rightarrow \infty$. We use the traditional *dipole approximation* which leads to the well-known *Thomson formula*.

For this purpose, let us expand the charge density in the Maxwell equations (11.13) in a Taylor series of the type (42.11):

$$(11.21) \quad e\delta(\mathbf{x} - \mathbf{x}(t)) = e\delta(\mathbf{x}) + e\mathbf{x}(t) \cdot \nabla \delta(\mathbf{x}) + \frac{1}{2}e(\mathbf{x}(t) \cdot \nabla)^2 \delta(\mathbf{x}) + \dots, \quad t > 0.$$

Here, the first term is static, and the corresponding Maxwell field is static with zero energy flux. The second term corresponds to the Hertzian dipole with dipole moment $\mathbf{p}(t) := e\mathbf{x}(t)$. The next terms of (11.21) give small contributions to the total energy radiation to infinity by (11.17).

Exercise 11.3 i) *Prove the convergence of the Taylor series (11.21) in the sense of (42.14) with $\psi \in \mathcal{H}_a(\mathbb{R}^3)$ for $a > \max |\mathbf{x}(t)| = 2 \frac{|e|A}{\mu k c^2}$.*

ii) *Prove that the energy flux $S_r(t, \mathbf{x})$ for large $|x|$ is determined by the second term of the expansion (11.21) up to an error $\mathcal{O}(\beta)/|x|^2$. **Hint:** Use the methods of Exercise 45.6.*

So we can use the Hertzian formula (45.72) for the dipole radiation (Section 45.6);

$$(11.22) \quad \mathbf{S}_r(t, \mathbf{x}) \approx \frac{\sin^2 \chi}{4\pi c^3 |\mathbf{x}|^2} \ddot{\mathbf{p}}^2(t - |\mathbf{x}|/c) \mathbf{n}, \quad |\mathbf{x}| \rightarrow \infty,$$

where χ is the angle between $\ddot{\mathbf{p}}(t - |\mathbf{x}|/c)$ and $\mathbf{n} := \mathbf{x}/|\mathbf{x}|$. By (11.18) and (11.5), we have

$$(11.23) \quad \ddot{\mathbf{p}}(t) = e\ddot{\mathbf{x}}(t) = \frac{e^2}{\mu} \mathbf{E}_0 \cos k(-ct).$$

Hence,

$$(11.24) \quad \ddot{\mathbf{p}}(t)^2 = \left(\frac{e^2}{\mu}\right)^2 \mathbf{E}_0^2 \cos^2 kct.$$

Definition 11.4 Denote by θ the angle between \mathbf{n} and \mathbf{e}_3 , and by φ the (azimuthal) angle between \mathbf{e}_1 and the plane $(\mathbf{n}, \mathbf{e}_3)$.

Exercise 11.5 Check that $\cos \chi = \cos \varphi \sin \theta$. **Solution:** $\cos \chi$ is the projection of \mathbf{n} onto the vector $\ddot{\mathbf{p}}(t - |\mathbf{x}|/c)$, which is parallel to \mathbf{e}_1 . The projection of \mathbf{n} onto the plane $\Pi = (\mathbf{e}_1, \mathbf{e}_2)$ has length $\sin \theta$. Finally, the angle between this projection and \mathbf{e}_1 equals φ .

Corollary 11.6 $\sin^2 \chi = 1 - \cos^2 \varphi \sin^2 \theta$.

Therefore,

$$(11.25) \quad \mathbf{S}_r(t, \mathbf{x}) \approx \frac{1 - \cos^2 \varphi \sin^2 \theta}{4\pi c^3 |\mathbf{x}|^2} \left(\frac{e^2}{\mu}\right)^2 \mathbf{E}_0^2 \cos^2 kc(t - |\mathbf{x}|/c) \mathbf{n}, \quad |\mathbf{x}| \rightarrow \infty.$$

Hence, for large $|\mathbf{x}|$, the energy flux is directed along \mathbf{n} . The corresponding intensity is obtained by the standard replacement of $\cos^2 kc(t - |\mathbf{x}|/c)$ by $1/2$:

$$(11.26) \quad \begin{aligned} I_r(\mathbf{x}) &:= \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_0^T \mathbf{S}_r(t, \mathbf{x}) dt \right| \approx \frac{1 - \cos^2 \varphi \sin^2 \theta}{8\pi |\mathbf{x}|^2} \left(\frac{e^2}{\mu c^2}\right)^2 \mathbf{E}_0^2 \\ &= \left(\frac{e^2}{\mu c^2}\right)^2 \frac{1 - \cos^2 \varphi \sin^2 \theta}{|\mathbf{x}|^2} I_0. \end{aligned}$$

Therefore, the *intensity per unit angle* $\mathcal{I}_r := I_r(\mathbf{x}) |\mathbf{x}|^2$ is

$$(11.27) \quad \mathcal{I}_r(\varphi, \theta) \approx \left(\frac{e^2}{\mu c^2}\right)^2 (1 - \cos^2 \varphi \sin^2 \theta) I_0.$$

Hence, the *differential cross-section* is given by the *Thomson formula*

$$(11.28) \quad D(\varphi, \theta) := \frac{\mathcal{I}_r(\varphi, \theta)}{I_0} \approx \left(\frac{e^2}{\mu c^2}\right)^2 (1 - \cos^2 \varphi \sin^2 \theta).$$

Finally, the *total cross-section* is given by

$$(11.29) \quad \begin{aligned} T &:= \int \mathcal{I}_r(\varphi, \theta) d\Omega / I_0 = \int D(\varphi, \theta) d\Omega \approx \left(\frac{e^2}{\mu c^2}\right)^2 \int (1 - \cos^2 \varphi \sin^2 \theta) d\Omega \\ &= \left(\frac{e^2}{\mu c^2}\right)^2 \left[4\pi - \int_0^{2\pi} \cos^2 \varphi d\varphi \int_0^\pi \sin^3 \theta d\theta \right] = \left(\frac{e^2}{\mu c^2}\right)^2 \frac{8\pi}{3}. \end{aligned}$$

Remark 11.7 The differential cross-section (11.28) depends on φ , hence it is not invariant with respect to rotations around \mathbf{e}_3 . This reflects the fact that the incident wave is linearly polarized. If we consider light with random polarization, then the differential cross-section is given by (11.28) with $1/2$ instead of $\cos^2 \varphi$.

12 Quantum Scattering of Light. Zero Order Approximation

We study scattering of light described as a plane wave by a quantum hydrogen atom in its ground state, i.e. in the state of the lowest possible energy. We derive an energy flux and a differential cross-section.

The ground state energy is $E_1 = -2\pi\hbar R = -\mu e^4/(2\hbar^2)$, and the corresponding wave function (8.34) is $\psi_1(\mathbf{x}) = C_1 e^{-|\mathbf{x}|/a_1}$ (we assume that the atom is situated at the origin). Then corresponding solution to the Schrödinger equation is

$$(12.1) \quad \psi_1(t, \mathbf{x}) = C_1 e^{-|\mathbf{x}|/a_1} e^{-i\omega_1 t}, \quad \omega_1 = E_1/\hbar = -\mu e^4/(2\hbar^3).$$

We want to describe the scattering of the plane wave (11.5) by the atom in the ground state. The scattering is described by the coupled Maxwell-Schrödinger equations in the Born approximation (7.27), (7.28). In the Lorentz gauge we have

$$(12.2) \quad [i\hbar\partial_t - e\phi^{\text{ext}}(\mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}^{\text{ext}}(t, \mathbf{x})]^2\psi(t, \mathbf{x}),$$

$$(12.3) \quad \begin{cases} \frac{1}{4\pi}\square\phi(t, \mathbf{x}) = \rho(t, \mathbf{x}) = e|\psi(t, \mathbf{x})|^2, \\ \frac{1}{4\pi}\square\mathbf{A}(t, \mathbf{x}) = \frac{\mathbf{j}(t, \mathbf{x})}{c} = \frac{e}{\mu c}[-i\hbar\nabla - \frac{e}{c}\mathbf{A}^{\text{ext}}(t, \mathbf{x})]\psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}), \end{cases}$$

where $\phi^{\text{ext}} = -e/|\mathbf{x}|$ is the Coulomb field of the nucleus and \mathbf{A}^{ext} stands for the incident wave (11.4) (cf. (11.11)),

$$(12.4) \quad \mathbf{A}^{\text{ext}}(t, \mathbf{x}) = A \sin k(x_3 - ct)\Theta(-\mathbf{x}_3 + ct)(1, 0, 0).$$

12.1 Atom Form Factor

At zero order approximation (in A), the wave function is unperturbed, $\psi(t, \mathbf{x}) = \psi_1(t, \mathbf{x})$. The corresponding approximation to the radiation fields ϕ, \mathbf{A} is given by the solutions to the Maxwell equations (12.3) with $\psi(t, \mathbf{x}) = \psi_1(t, \mathbf{x})$. Then the charge density $\rho(t, \mathbf{x})$ is static and the corresponding potential ϕ is static with zero energy flux. Therefore, it suffices to solve the equation for \mathbf{A} with the current

$$(12.5) \quad \frac{\mathbf{j}(t, \mathbf{x})}{c} = \frac{e}{\mu c}[-i\hbar\nabla\psi_1(t, \mathbf{x}) \cdot \psi_1(t, \mathbf{x}) - \frac{e}{c}\mathbf{A}^{\text{ext}}(t, \mathbf{x})\psi_1(t, \mathbf{x}) \cdot \psi_1(t, \mathbf{x})].$$

Here, the first term on the RHS is zero since the corresponding eigenfunction $e^{-|\mathbf{x}|/a_1}$ is real:

$$(12.6) \quad i\hbar\nabla\psi_1(t, \mathbf{x}) \cdot \psi_1(t, \mathbf{x}) := \text{Re}\left(i\hbar\nabla\psi_1(t, \mathbf{x})\bar{\psi}_1(t, \mathbf{x})\right) = |C|^2\text{Re}\left(i\hbar(\nabla e^{-|\mathbf{x}|/a_1})e^{-|\mathbf{x}|/a_1}\right) = 0.$$

Therefore, the current is reduced to

$$(12.7) \quad \frac{\mathbf{j}(t, \mathbf{x})}{c} = -\frac{e^2}{\mu c^2}\mathbf{A}^{\text{ext}}(t, \mathbf{x})|\psi_1(\mathbf{x})|^2.$$

Let us split the solution

$$(12.8) \quad \mathbf{A}(t, \mathbf{x}) = \mathbf{A}^{\text{ext}}(t, \mathbf{x}) + \mathbf{A}_r(t, \mathbf{x}),$$

where $\mathbf{A}_r(t, \mathbf{x})$ is the *radiated field*. Then (12.3) becomes

$$(12.9) \quad \begin{aligned} \square\mathbf{A}_r(t, \mathbf{x}) &= -4\pi\frac{e^2}{\mu c^2}A \sin k(x_3 - ct)|\psi_1(\mathbf{x})|^2\Theta(-\mathbf{x}_3 + ct)(1, 0, 0) \\ &= -4\pi\frac{e^2}{\mu c^2}\text{Im}Ae^{ik(x_3 - ct)}|\psi_1(\mathbf{x})|^2\Theta(-\mathbf{x}_3 + ct)(1, 0, 0) =: f(t, \mathbf{x}) \end{aligned}$$

since \mathbf{A}^{ext} is a solution to the homogeneous equation.

The radiation field could be characterized uniquely by the initial condition of type (11.14) if the atom radius would be zero as in the classical case of previous lecture. In our case the radiation field is of finite energy, i.e.

$$(12.10) \quad E_r(t) := \int_{\mathbb{R}^3} \left(|\dot{\mathbf{A}}_r(t, \mathbf{x})|^2 + (|\nabla \mathbf{A}_r(t, \mathbf{x})|^2) \right) dx < \infty, \quad t \in \mathbb{R}.$$

An exact characterization of the radiation field would be given by the limit $E_r(t) \rightarrow 0$, $t \rightarrow -\infty$, however (12.10) is sufficient for our purposes. Namely, let us demonstrate that Theorem 43.3 and (12.10) with $t = 0$ imply the *limiting amplitude principle* holds, i.e.

$$(12.11) \quad \mathbf{A}_r(t, \mathbf{x}) \sim \text{Im } A_r(\mathbf{x}) e^{-ikct}, \quad t \rightarrow \infty.$$

In particular, the zero order approximation to the radiation field has the same frequency as the incident wave.

Indeed, the asymptotics (12.11) follow similarly to (43.30). That is, Theorem 43.3 implies that the long-time asymptotics of the finite-energy solution $\mathbf{A}_r(t, \mathbf{x})$ is given by the retarded potential

$$(12.12) \quad \begin{aligned} \mathbf{A}_r(t, \mathbf{x}) &\sim \int \frac{f(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}}{4\pi |\mathbf{x} - \mathbf{y}|} \\ &= -\frac{e^2}{\mu c^2} A \text{Im} \int \frac{e^{ik(y_3 - c(t - |\mathbf{x} - \mathbf{y}|/c))} |\psi_1(\mathbf{y})|^2 d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|} (1, 0, 0) \\ &= -\frac{e^2}{\mu c^2} A \text{Im} e^{-ikct} \int \frac{e^{ik(y_3 + |\mathbf{x} - \mathbf{y}|)} |\psi_1(\mathbf{y})|^2 d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|} (1, 0, 0). \end{aligned}$$

Let us find the asymptotics of the integral as $|\mathbf{x}| \rightarrow \infty$. For any fixed $\mathbf{y} \in \mathbb{R}^3$,

$$(12.13) \quad |\mathbf{x} - \mathbf{y}| = |\mathbf{x}| - \mathbf{y} \cdot \mathbf{n} + o(1), \quad |\mathbf{x}| \rightarrow \infty,$$

where $\mathbf{n} = \mathbf{n}(x) = \mathbf{x}/|\mathbf{x}|$. Let us also write $\mathbf{y}_3 = \mathbf{y} \cdot \mathbf{e}_3$. Then we get

$$(12.14) \quad \mathbf{A}_r(t, \mathbf{x}) \sim -\frac{e^2}{\mu c^2} A \text{Im} e^{-ikct} \frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|} \int e^{ik\mathbf{y} \cdot (\mathbf{e}_3 - \mathbf{n})} |\psi_1(\mathbf{y})|^2 d\mathbf{y} (1, 0, 0), \quad |\mathbf{x}| \rightarrow \infty.$$

Next, we evaluate the last integral. Set $K := k|\mathbf{e}_3 - \mathbf{n}|$ and denote by θ the angle between \mathbf{n} and \mathbf{e}_3 . Then

$$(12.15) \quad K = K(k, \theta) = k\sqrt{n_1^2 + n_2^2 + (1 - n_3)^2} = k\sqrt{2(1 - n_3)} = k\sqrt{2(1 - \cos \theta)} = 2k \sin \frac{\theta}{2}.$$

Denote by α the angle between \mathbf{y} and $\mathbf{e}_3 - \mathbf{n}$, and by φ the azimuthal angle around $\mathbf{e}_3 - \mathbf{n}$. Finally, let us take into account that the ground state $\psi_1(\mathbf{y}) = \psi_1^r(|\mathbf{y}|)$ is spherically symmetric. Then the integral becomes

$$(12.16) \quad \begin{aligned} &\int_0^\infty |\mathbf{y}|^2 d|\mathbf{y}| \int_0^\pi \sin \alpha d\alpha \int_0^{2\pi} d\varphi e^{iK \cos \alpha |\mathbf{y}|} |\psi_1(\mathbf{y})|^2 \\ &= 4\pi \int_0^\infty \frac{\sin K|\mathbf{y}|}{K|\mathbf{y}|} |\psi_1^r(|\mathbf{y}|)|^2 |\mathbf{y}|^2 d|\mathbf{y}| =: F_a(k, \theta) \end{aligned}$$

which is called the *atom form factor*.

Exercise 12.1 Calculate the last integral.

Since $F_a(\theta)$ is real, the asymptotics (12.14) becomes

$$(12.17) \quad \mathbf{A}_r(t, \mathbf{x}) \sim -\frac{e^2}{\mu c^2} A \frac{\sin k(|\mathbf{x}| - ct)}{|\mathbf{x}|} F_a(k, \theta) (1, 0, 0), \quad |\mathbf{x}| \rightarrow \infty.$$

12.2 Energy flux

We still have to calculate the Maxwell field and the Poynting vector corresponding to this vector potential. It suffices to compare the expressions (12.17) with the vector potential of the Hertzian dipole (formula (45.67) of Exercise 45.6):

$$(12.18) \quad \mathbf{A}(t, \mathbf{x}) = \frac{1}{c} \frac{\dot{\mathbf{p}}(t - r/c)}{r}.$$

This is identical to (12.17) if $F_a(\theta) = 1$ and

$$(12.19) \quad \mathbf{p}(t) := \frac{e^2}{\mu c^2 k} A \cos kct \mathbf{e}_3.$$

Therefore, the energy flux $\mathbf{S}(t, \mathbf{x})$ corresponding to (12.17) is given, up to an error $\mathcal{O}(|x|^{-3})$, by the Hertzian formula (11.22), with a factor $|F_a(\theta)|^2$. This follows from the fact that any differentiation of the form factor $F_a(\theta)$ in \mathbf{x}_k gives an additional factor with the decay $\mathcal{O}(|x|^{-1})$ since the form factor is a homogeneous function of \mathbf{x} . Finally, for the function (12.19), $\ddot{\mathbf{p}}(t)$ coincides with (11.23). Therefore, (11.25) gives in our case

$$(12.20) \quad \mathbf{S}(t, \mathbf{x}) \sim |F_a(\theta)|^2 \frac{1 - \cos^2 \varphi \sin^2 \theta}{4\pi c^3 |\mathbf{x}|^2} \left(\frac{e^2}{\mu}\right)^2 E_0^2 \cos^2 kc(t - |\mathbf{x}|/c) \mathbf{n}, \quad |\mathbf{x}| \rightarrow \infty.$$

Then the intensity and differential cross-section in our case also contain the additional factor $|F_a(\theta)|^2$. Hence, the differential cross-section coincides with the Thomson formula (11.28) up to the atom form factor:

$$(12.21) \quad D(k, \varphi, \theta) = |F_a(k, \theta)|^2 \left(\frac{e^2}{\mu c^2}\right)^2 (1 - \cos^2 \varphi \sin^2 \theta).$$

13 Light Scattering at Small Frequencies: Short-Range Scattering

Now we take into account the change of the ground state induced by the incident light (11.4), in the first order with respect to A , for small frequencies of light. This change describes the *polarization* of the hydrogen atom (see previous lecture). It defines the corresponding *dispersion* and the *combinational scattering*.

13.1 First Order Approximation to the Ground State

To calculate the first approximation to the ground state, let us use the Schrödinger equation (12.2): for small amplitudes $|A|$,

$$(13.1) \quad \psi(t, \mathbf{x}) = \psi_1(t, \mathbf{x}) + Aw(t, \mathbf{x}) + \mathcal{O}(A^2).$$

Substituting into (12.2), we obtain in first order in A ,

$$(13.2) \quad A[i\hbar\partial_t - e\phi^{\text{ext}}(\mathbf{x})]w(t, \mathbf{x}) = A\frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}}]^2w(t, \mathbf{x}) + \frac{i\hbar e}{\mu c}\mathbf{A}^{\text{ext}}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}}\psi_1(t, \mathbf{x}),$$

since $\psi_1(t, \mathbf{x})$ is a solution to Equation (12.2) with $\mathbf{A}^{\text{ext}} = 0$ and the Coulomb potential $\phi^{\text{ext}} = -e/|\mathbf{x}|$ of the nucleus. By (11.4) and (12.1), we have the following long-time asymptotics for the RHS of (13.2),

$$(13.3) \quad \begin{aligned} \frac{i\hbar e}{\mu c}\mathbf{A}^{\text{ext}}(t, \mathbf{x}) \cdot \nabla_{\mathbf{x}}\psi_1(t, \mathbf{x}) &\sim \frac{i\hbar e}{\mu c}A \sin k(x_3 - ct)(1, 0, 0) \cdot \nabla_{\mathbf{x}}\psi_1(\mathbf{x})e^{-i\omega_1 t} \\ &= \frac{A\hbar e}{2\mu c}[e^{ik(x_3 - ct)} - e^{-ik(x_3 - ct)}]e^{-i\omega_1 t}\nabla_1\psi_1(\mathbf{x}) \\ &= \psi_+(\mathbf{x})e^{-i(\omega_1 + \omega)t} - \psi_-(\mathbf{x})e^{-i(\omega_1 - \omega)t}, \end{aligned}$$

where $\omega := kc$. Now let us apply the *limiting amplitude principle* (43.28):

$$(13.4) \quad w(t, \mathbf{x}) = w_+(\mathbf{x})e^{-i(\omega_1 + \omega)t} - w_-(\mathbf{x})e^{-i(\omega_1 - \omega)t} + \sum_l C_l \psi_l(\mathbf{x})e^{-i\omega_l t} + r(t, \mathbf{x}),$$

where $w_{\pm}(\mathbf{x})$ are the *limiting amplitudes*, $\psi_l(\mathbf{x})$ stand for the eigenfunctions of the discrete spectrum of the homogeneous Schrödinger equation $[i\hbar\partial_t - e\phi^{\text{ext}}(\mathbf{x})]w(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}}]^2w(t, \mathbf{x})$, which corresponds to (13.2), and $r(t, \mathbf{x}) \rightarrow 0$, $t \rightarrow \infty$, in an appropriate norm. Below, we will omit the sum over the discrete spectrum in the RHS of (13.4). According to Remark 43.11,

- i) the *limiting amplitudes* $w_{\pm}(\mathbf{x})$ characterize the response of the atom to the incident wave and depend on the frequency ω and Amplitude A of the incident wave, and do not depend on initial data.
- ii) The sum over the discrete spectrum on the RHS of (13.4) depends only on initial data and does not depend on the incident wave.

Hence, the contribution of the incident wave to the electric current and magnetization (see below), in principle, can be singled out in an experimental observation of the considered scattering problem. Finally, we identify

$$(13.5) \quad w(t, \mathbf{x}) \sim w_+(\mathbf{x})e^{-i(\omega_1 + \omega)t} - w_-(\mathbf{x})e^{-i(\omega_1 - \omega)t}, \quad t \rightarrow +\infty.$$

For w_{\pm} we get the equations (cf. (43.23))

$$(13.6) \quad [\hbar(\omega_1 \pm \omega) - e\phi^{\text{ext}}(\mathbf{x})]w_{\pm}(\mathbf{x}) - \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}}]^2w_{\pm}(\mathbf{x}) = \psi_{\pm}(\mathbf{x}) := \frac{\hbar e}{2\mu c}e^{\pm ikx_3}\nabla_1\psi_1(\mathbf{x}).$$

Now we introduce the following condition:

$$(13.7) \quad \text{spectral bound:} \quad |\omega| < |\omega_1|.$$

Remark 13.1 Let us note that for the case of the hydrogen atom we have $|\omega_1| = \frac{\mu e^4}{2\hbar^3} \approx 20,5 \cdot 10^{15} \text{rad/sec}$. Hence, the bound (13.7) holds for the frequencies $|\omega| < 3,27 \cdot 10^{15} \text{Hz}$ or wave numbers $k < |\omega_1|/c \approx 68 \cdot 10^7 \text{ m}^{-1}$ and wave lengths $\lambda > 0,91176 \cdot 10^{-5} \text{ cm} = 911,76 \text{ \AA}$.

The frequency bound implies that $\omega_1 \pm \omega < 0$, hence the values $\hbar(\omega_1 \pm \omega)$ do not belong to the continuous spectrum of the stationary Schrödinger equation (13.6). We will also assume that $\hbar(\omega_1 \pm \omega)$ do not belong to the discrete spectrum. The last assumption is unessential since the coincidence $\hbar(\omega_1 \pm \omega)$ with an eigenvalue $\hbar\omega_n$ is an event of codimension one, i.e. probability zero. Therefore, the solutions decay at infinity in the following sense

$$(13.8) \quad w_{\pm} \in L^2(\mathbb{R}^3)$$

(see Remark 43.10 ii)). For a moment let us neglect the external potential $\phi^{\text{ext}}(\mathbf{x})$ in the equation (13.6). Then it can be rewritten in the form

$$(13.9) \quad [z + \nabla_{\mathbf{x}}^2]w(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3,$$

where $z < 0$ and $|f(\mathbf{x})| \leq Ce^{-\varepsilon|\mathbf{x}|}$ with an $\varepsilon > 0$.

Exercise 13.2 Prove that $w_{\pm}(\mathbf{x})$ decay exponentially at infinity, like $\psi_{\pm}(\mathbf{x})$:

$$(13.10) \quad |w_{\pm}(\mathbf{x})| \leq Ce^{-\varepsilon_1|\mathbf{x}|}, \quad \mathbf{x} \in \mathbb{R}^3,$$

where $\varepsilon_1 > 0$. **Hint:** Apply the Fourier transform to prove that i) the solution is unique in the class of tempered distributions and ii) the solution is a convolution, $w = f * E$, where $E(\mathbf{x})$ is the fundamental solution $E(\mathbf{x}) = -e^{-\kappa|\mathbf{x}|}/(4\pi|\mathbf{x}|)$ with $\kappa := \sqrt{-z} > 0$. Then $\varepsilon_1 = \min(\varepsilon, \kappa) > 0$.

Remark 13.3 We have supposed that the atom is in its groundstate, ψ_1 , in the remote past, $t \rightarrow -\infty$. The representation (13.1), the asymptotics (13.4) and (13.8) imply that, roughly speaking, the light scattering modify the groundstate in the first order in A , in the long-time limit $t \rightarrow \infty$. The condition (13.8) means that the scattering process is essentially concentrated in a bounded region of space. Hence, the process corresponds to a *short-range scattering*.

Further, let us calculate the modified groundstate using the spectral resolution of the Schrödinger operator in the equation (13.6). First, let us expand the RHS,

$$(13.11) \quad \frac{\hbar e}{2\mu c} e^{\pm ikx_3} \nabla_1 \psi_1(\mathbf{x}) = \sum_l a_l^{\pm} \psi_l(\mathbf{x}),$$

where \sum_l stands for the sum over the discrete spectrum and the integral over the continuous spectrum. Then the solutions w_{\pm} have the form

$$(13.12) \quad w_{\pm}(\mathbf{x}) = \sum_l \frac{a_l^{\pm} \psi_l(\mathbf{x})}{\hbar(\omega_1 \pm \omega - \omega_l)}.$$

Therefore, (13.5) becomes

$$(13.13) \quad w(t, \mathbf{x}) \sim \sum_l \frac{a_l^+ \psi_l(\mathbf{x})}{\hbar(\omega_1 + \omega - \omega_l)} e^{-i(\omega_1 + \omega)t} - \sum_l \frac{a_l^- \psi_l(\mathbf{x})}{\hbar(\omega_1 - \omega - \omega_l)} e^{-i(\omega_1 - \omega)t}.$$

Let us calculate the coefficients a_l^{\pm} . Formally,

$$(13.14) \quad a_l^{\pm} = \frac{\hbar e}{2\mu c} \int \bar{\psi}_l(\mathbf{x}) e^{\pm ikx_3} \nabla_1 \psi_1(\mathbf{x}) d\mathbf{x}.$$

Let us assume that $k \ll 1$ and set $e^{\pm ikx_3} = 1$. Then a_l^\pm are approximately identical, and by partial integration,

$$\begin{aligned}
a_l^\pm &\approx \frac{\hbar e}{4\mu c} \int [\bar{\psi}_l(\mathbf{x}) \nabla_1 \psi_1(\mathbf{x}) - \nabla_1 \bar{\psi}_l(\mathbf{x}) \psi_1(\mathbf{x})] d\mathbf{x} \\
&= \frac{\hbar e}{4\mu c} e^{i(\omega_1 - \omega_l)t} \int [\bar{\psi}_l(t, \mathbf{x}) \nabla_1 \psi_1(t, \mathbf{x}) - \nabla_1 \bar{\psi}_l(t, \mathbf{x}) \psi_1(t, \mathbf{x})] d\mathbf{x} \\
(13.15) \quad &= i \frac{1}{2c} e^{i(\omega_1 - \omega_l)t} \int \mathbf{j}_{1l}^1(t, \mathbf{x}) d\mathbf{x}.
\end{aligned}$$

Here $\psi_l(t, \mathbf{x}) := e^{-i\omega_l t} \psi_l(\mathbf{x})$ and $\mathbf{j}_{1l}^1(t, \mathbf{x})$ is the first component of the current (7.23) corresponding to $\psi_1(t, \mathbf{x})$ and $\bar{\psi}_l(t, \mathbf{x})$, since the functions are solutions to the Schrödinger equation in (7.4) with $A(t, x) + A^{\text{ext}}(t, x) = 0$. Then the identity (10.20) implies,

$$(13.16) \quad a_l^\pm \approx \frac{\omega_{1l}}{2c} \int \mathbf{x}^1 e \psi_1(\mathbf{x}) \bar{\psi}_l(\mathbf{x}) d\mathbf{x} = \frac{e\omega_{1l}}{2c} \mathbf{x}_{1l}^1 =: a_l,$$

where $\omega_{1l} := \omega_1 - \omega_l$ and $\mathbf{x}_{1l}^1 := \int \mathbf{x}^1 \psi_1(\mathbf{x}) \bar{\psi}_l(\mathbf{x}) d\mathbf{x}$. Finally, (13.13) becomes,

$$(13.17) \quad w(t, \mathbf{x}) \sim \sum_l a_l \psi_l(\mathbf{x}) \left(\frac{e^{-i\omega t}}{\hbar(\omega_{1l} + \omega)} - \frac{e^{i\omega t}}{\hbar(\omega_{1l} - \omega)} \right) e^{-i\omega_1 t},$$

and the wave function (13.1) reads

$$\begin{aligned}
\psi(t, \mathbf{x}) &= \left[\psi_1(\mathbf{x}) + A \sum_l a_l \psi_l(\mathbf{x}) \left(\frac{e^{-i\omega t}}{\hbar(\omega_{1l} + \omega)} - \frac{e^{i\omega t}}{\hbar(\omega_{1l} - \omega)} \right) \right] e^{-i\omega_1 t} + \mathcal{O}(A^2) \\
(13.18) \quad &= \left[\psi_1(\mathbf{x}) + A \Sigma(t, \mathbf{x}) \right] e^{-i\omega_1 t} + \mathcal{O}(A^2).
\end{aligned}$$

13.2 Polarization and Dispersion: Kramers-Kronig Formula

Let us calculate the corresponding electric dipole moment. First, the charge density is given by

$$\begin{aligned}
\rho(t, \mathbf{x}) &= e\psi(t, \mathbf{x}) \bar{\psi}(t, \mathbf{x}) = e(\psi_1(\mathbf{x}) + A \Sigma(t, \mathbf{x})) (\bar{\psi}_1(\mathbf{x}) + A \bar{\Sigma}(t, \mathbf{x})) + \mathcal{O}(A^2) \\
(13.19) \quad &= e|\psi_1(t, \mathbf{x})|^2 + eA \left[\Sigma^+ e^{i\omega t} + \Sigma^- e^{-i\omega t} \right] + \mathcal{O}(A^2),
\end{aligned}$$

where

$$(13.20) \quad \Sigma^+ = \sum_l \left(\frac{\psi_1 \bar{a}_l \bar{\psi}_l}{\hbar(\omega_{1l} + \omega)} - \frac{a_l \psi_l \bar{\psi}_1}{\hbar(\omega_{1l} - \omega)} \right), \quad \Sigma^- = \bar{\Sigma}^+.$$

Therefore, the electric dipole moment equals, mod $\mathcal{O}(A^2)$,

$$(13.21) \quad \mathbf{p}(t) := \int \mathbf{x} \rho(t, \mathbf{x}) d\mathbf{x} = \mathbf{p}_{11} + \mathbf{P}(t),$$

where $\mathbf{p}_{11} := \int \mathbf{x} e |\psi_1(t, \mathbf{x})|^2 d\mathbf{x} = 0$ by spherical symmetry, and

$$(13.22) \quad \mathbf{P}(t) = eA \left[\sum_l \left(\frac{\bar{a}_l \mathbf{x}_{1l}}{\hbar(\omega_{1l} - \omega)} - \frac{a_l \bar{\mathbf{x}}_{1l}}{\hbar(\omega_{1l} + \omega)} \right) e^{i\omega t} + \sum_l \left(\frac{a_l \bar{\mathbf{x}}_{1l}}{\hbar(\omega_{1l} - \omega)} - \frac{\bar{a}_l \mathbf{x}_{1l}}{\hbar(\omega_{1l} + \omega)} \right) e^{-i\omega t} \right],$$

where $\mathbf{x}_{1l} := \int \mathbf{x} \psi_1 \bar{\psi}_l d\mathbf{x}$. By symmetry arguments, the vector $\mathbf{P}(t)$ is directed along \mathbf{e}_1 . Indeed, the invariance of $\mathbf{p}(t)$ with respect to the reflection $\mathbf{x}_2 \mapsto -\mathbf{x}_2$ is obvious. The invariance with respect

to the reflection $\mathbf{x}_3 \mapsto -\mathbf{x}_3$ follows from (13.11)–(13.13) since we set $k = 0$. Therefore, substituting $a_l = e\omega_{1l}\mathbf{x}_{1l}^1/(2c)$ and projecting \mathbf{x}_{1l} onto \mathbf{e}_1 , we get

$$(13.23) \quad \mathbf{p}(t) = A\mathbf{e}_1 \frac{4ke^2}{\hbar} \sum_l \frac{\omega_{1l}|\mathbf{x}_{1l}^1|^2}{\omega_{1l}^2 - \omega^2} \cos \omega t$$

At last, let us average this expression with respect to all orientations of the atom. Then we obtain

$$(13.24) \quad \overline{\mathbf{p}}(t) := A\mathbf{e}_1 \frac{4ke^2}{\hbar} \sum_l \frac{\omega_{1l}\overline{|\mathbf{x}_{1l}^1|^2}}{\omega_{1l}^2 - \omega^2} \cos \omega t,$$

since the average of \mathbf{p}_{11} is obviously zero. Let us calculate the permittivity (42.30) of the atomic hydrogen in the ground state ψ_1 . We denote by $\mathbf{E}(t)$ the electric field at the position $\mathbf{x} = 0$ of the atom: by (11.5), we have $\mathbf{E}(t) = kA \cos \omega t \mathbf{e}_1$.

Corollary 13.4 *The permittivity of the atomic hydrogen in the ground state is given by the Kramers-Kronig formula*

$$(13.25) \quad \chi_e = n|\overline{\mathbf{p}}(t)|/|\mathbf{E}(t)| = n \frac{4e^2}{\hbar} \sum_l \frac{\omega_{1l}\overline{|\mathbf{x}_{1l}^1|^2}}{\omega_{1l}^2 - \omega^2},$$

where n is the number of atoms per unit volume.

This formula has the same analytic structure as its classical analog (45.80) of Exercise 45.7. This allows us to express the hydrogen electric susceptibility $\varepsilon = 1 + 4\pi\chi$ (see (42.31)) and hence the refraction coefficient $n \sim \sqrt{\varepsilon}$ in the case of the magnetic susceptibility $\mu \sim 1$ (see (45.99)).

13.3 Combinational Scattering

Formula (13.18) gives the first order correction to the unperturbed ground state $\psi_1(\mathbf{x})e^{-i\omega_1 t}$ in the presence of the incident wave (11.5). Similarly, for any unperturbed stationary state $\psi_j(\mathbf{x})e^{-i\omega_j t}$, the corresponding first order correction is obtained by changing the index from 1 to j . Therefore, the first order correction to a general unperturbed solution of the type (10.2) reads

$$(13.26) \quad \psi(t, \mathbf{x}) = \sum_j c_j(T) \left[\psi_j(\mathbf{x}) + A \sum_l a_l \psi_l(\mathbf{x}) \left(\frac{e^{-i\omega t}}{\hbar(\omega_{jl} + \omega)} - \frac{e^{i\omega t}}{\hbar(\omega_{jl} - \omega)} \right) \right] e^{-i\omega_j t}.$$

Hence, we can apply the methods of Section 10 to calculate the spectrum of the dipole radiation of the hydrogen atom in the presence of light. Now we get a new set of frequencies: in the first order approximation in A , the spectrum is contained in the set $\{\omega_{jj'}, \omega_{jj'} \pm \omega\}$. The corresponding selection rules also can be obtained by the methods of Section 10.

14 Light Scattering in Continuous Spectrum: Photoeffect

For light with a frequency which is larger than the binding energy of the stationary electron state, scattering of the electron into the continuous spectrum may occur. Within the coupled Maxwell-Schrödinger system this scattering is described by a weakly decaying contribution to the electron wave function, which corresponds to a non-zero electron current radiating to infinity.

14.1 Radiation in Continuous Spectrum

Let us consider the scattering of light with frequency

$$(14.1) \quad \omega > |\omega_1|.$$

Then $\omega_1 - \omega < 0$ and $\omega_1 + \omega = \omega - |\omega_1| > 0$. This means that $\hbar(\omega_1 + \omega)$ belongs to the continuous spectrum of the stationary Schrödinger equation (13.6) which implies a slow decay of the wave function $|w_+(\mathbf{x})| \sim 1/|\mathbf{x}|$ at $|\mathbf{x}| \rightarrow \infty$. We will calculate the long-range asymptotics of $w_+(\mathbf{x})$ and obtain the main term of the radiation in the form

$$(14.2) \quad Aw_+(\mathbf{x})e^{-i(\omega_1+\omega)t} \sim A \frac{a(\varphi, \theta)}{|\mathbf{x}|} e^{i[k_r|\mathbf{x}| - (\omega_1+\omega)t]}, \quad |\mathbf{x}| \rightarrow \infty.$$

On the other hand, $\hbar(\omega_1 - \omega)$ does not belong to the continuous spectrum. Let us also assume that $\hbar(\omega_1 - \omega)$ does not belong to the discrete spectrum, hence

$$(14.3) \quad w_- \in L^2$$

similarly to (13.8). We will deduce from the asymptotics (14.2) the following stationary electric current at infinity,

$$(14.4) \quad \mathbf{j}(t, \mathbf{x}) \sim A^2 \frac{e\hbar k_r}{\mu} \frac{a^2(\varphi, \theta)}{|\mathbf{x}|^2} \mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \rightarrow \infty,$$

where $\mathbf{n}(\mathbf{x}) := \mathbf{x}/|\mathbf{x}|$. For the first time the formula was derived by Wentzel in 1927 ([100]) with the amplitude

$$(14.5) \quad a(\varphi, \theta) = C \sin \theta \cos \varphi.$$

Remark 14.1 The formula (14.4) describes the radiation of a non-zero electric current from the atom to infinity. Namely, The *slow decay* of the current (14.4) at $|\mathbf{x}| \rightarrow \infty$ implies that the *total electric current to infinity* does not vanish, i.e.

$$(14.6) \quad J_\infty := \lim_{R \rightarrow \infty} \int_{|\mathbf{x}|=R} \mathbf{j}(t, \mathbf{x}) dS(\mathbf{x}) \neq 0.$$

This corresponds to the *long-range scattering* under the condition (14.1). If the condition fails, the *fast decay* of type (13.10) implies that in this case the total electric current to infinity vanishes that corresponds to the *short-range scattering* under the condition (13.7) (see Remark 13.3).

Comment 14.2 The stationary nonvanishing current (14.6) formally contradicts the charge conservation for the atom. The contradiction is provided by the perturbation strategy which leaves the current (12.5) unchanged. To maintain the stationary photocurrent, one needs an external source (galvanic element, etc) to reimburse the charge decay.

Corollary 14.3 i) The Schrödinger theory explains the “red bound” of the photoeffect, $\omega > \omega_{\text{red}}$, and provides the value $\omega_{\text{red}} = |\omega_1|$.

ii) Similarly to (7.10), the radiated wave function (14.2) can be identified with a beam of *photoelectrons* with the density

$$(14.7) \quad d(\mathbf{x}) = \rho(t, \mathbf{x})/e = |Aw_+(\mathbf{x})|^2 = A^2 \frac{a^2(\varphi, \theta)}{|\mathbf{x}|^2}$$

and the velocities

$$(14.8) \quad \mathbf{v}(\mathbf{x}) = \mathbf{j}(t, \mathbf{x})/(ed(\mathbf{x})) = \frac{\hbar k_r}{\mu} \mathbf{n}(\mathbf{x}).$$

iii) At large $|\mathbf{x}|$, the radiated wave function (14.2) locally is close to the plane wave (15.21). Hence, the energy of one photoelectron is given by the formula of type (7.12). Hence, (14.2) implies that the energy of one photoelectron is given by the *Einstein formula*

$$(14.9) \quad E = \hbar(\omega - |\omega_1|).$$

14.2 Limiting Amplitude

Let us calculate the limiting amplitude $w_+(\mathbf{x})$. First, we rewrite Equation (13.6) in the form

$$(14.10) \quad [\nabla_{\mathbf{x}}^2 + k_r^2(\omega)]w_+(\mathbf{x}) = \frac{e}{\hbar c} e^{ikx_3} \nabla_1 \psi_1(\mathbf{x}) - \frac{2e^2 \mu}{\hbar^2 |\mathbf{x}|} w_+(\mathbf{x}),$$

where $k_r(\omega) = \sqrt{2\mu(\omega_1 + \omega)}/\hbar > 0$. For a moment, let us neglect the last term on the RHS. So, we get the Helmholtz equation

$$(14.11) \quad [\nabla_{\mathbf{x}}^2 + k_r^2(\omega)]w_+(\mathbf{x}) = f_+(\mathbf{x}) := \frac{e}{\hbar c} e^{ikx_3} \nabla_1 \psi_1(\mathbf{x}).$$

The limiting amplitude is given by the convolution (cf. (43.27))

$$(14.12) \quad w_+(\mathbf{x}) = - \int \frac{e^{ik_r(\omega)|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} f_+(\mathbf{y}) d\mathbf{y}.$$

This follows from the limiting absorption principle (43.32) since the fundamental solution $\frac{e^{ik_r(\omega+i\varepsilon)|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$ is a tempered distribution for small $\varepsilon > 0$. Indeed, $\text{Im } k_r(\omega + i\varepsilon) > 0$ for the fixed branch $k_r(\omega) > 0$.

Now we can calculate the asymptotics (14.2). Let us substitute the expression (14.11) for f_+ into (14.12). Then, after partial integration, we get

$$(14.13) \quad \begin{aligned} w_+(\mathbf{x}) &= \frac{e}{\hbar c} \int \nabla_1 \frac{e^{ik_r|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} e^{iky_3} \psi_1(\mathbf{y}) d\mathbf{y} \\ &= -\frac{ik_r e}{\hbar c} \int \frac{e^{ik_r|\mathbf{x}-\mathbf{y}|} (\mathbf{x}_1 - \mathbf{y}_1)}{4\pi|\mathbf{x}-\mathbf{y}|^2} e^{iky_3} \psi_1(\mathbf{y}) d\mathbf{y} + \mathcal{O}(|\mathbf{x}-\mathbf{y}|^{-2}). \end{aligned}$$

Substituting here $\mathbf{x}_1 - \mathbf{y}_1 = \sin \theta \cos \varphi |\mathbf{x} - \mathbf{y}|$, we get (14.2) with the amplitude (14.5) since the ground state $\psi_1(\mathbf{y})$ decays rapidly at infinity.

14.3 Electric Current at Infinity

Let us deduce (14.4) from (14.2). In the Born approximation the current is given by (7.28),

$$(14.14) \quad \mathbf{j} := -\frac{e}{\mu} [i\hbar \nabla \psi(t, \mathbf{x})] \cdot \psi(t, \mathbf{x}) - \frac{e^2}{\mu c} \mathbf{A}^{\text{ext}}(t, \mathbf{x}) |\psi(t, \mathbf{x})|^2.$$

The function w_- decays at infinity by (14.3). Therefore, (13.1) becomes

$$(14.15) \quad \psi(t, \mathbf{x}) \sim \psi_1(t, \mathbf{x}) + Aw_+(\mathbf{x})e^{-i(\omega_1+\omega)t}.$$

The contribution of the function ψ_1 to the current at infinity also is negligible due to the exponential decay at infinity. Therefore, (14.4) follows by substituting the asymptotics (14.2) into the first term on the RHS.

14.4 Coulomb Potential

Now let us discuss the general equation (14.10) instead of (14.11). Then (14.12) changes to

$$(14.16) \quad w_+(\mathbf{x}) = \int G_{k_r}(\mathbf{x}, \mathbf{y}) f_+(\mathbf{y}) d\mathbf{y},$$

where G_{k_r} is the corresponding Green function. For each fixed $\mathbf{y} \in \mathbb{R}^3$ the following asymptotics holds (cf. [89, Vol. II, formula (II.7.33)]),

$$(14.17) \quad G_{k_r}(\mathbf{x}, \mathbf{y}) \sim C_1 e^{i\gamma \log |\mathbf{x}-\mathbf{y}|} \frac{e^{ik_r |\mathbf{x}-\mathbf{y}|}}{4\pi |\mathbf{x}-\mathbf{y}|},$$

where $\gamma = \gamma(k_r) \in \mathbb{R}$. Therefore, the asymptotics (14.2), (14.4) follow by the same arguments as above, with the amplitude of the form (14.5).

Exercise 14.4 Deduce (14.2), (14.4), (14.5) from (14.16) and (14.17).

14.5 Shift in Angular Distribution: Sommerfeld Formula

The Wentzel calculation takes into account the first order interaction of the Maxwell and Schrödinger fields. Next, second order correction, has been obtained by Sommerfeld and Shur [90]. The corresponding corrected formula reads (see [89, Vol. II])

$$(14.18) \quad \mathbf{j}(t, \mathbf{x}) \sim \frac{\sin^2 \theta \cos^2 \varphi (1 + 4\beta \cos \theta)}{|\mathbf{x}|^2} \mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \rightarrow \infty.$$

Here $\beta = \frac{v}{c}$, where v is the velocity of the photoelectrons. The formula means an *increment* of the scattering amplitude for the angles $0 < \theta < \frac{\pi}{2}$ and a *decrement* of the scattering amplitude for the angles $\frac{\pi}{2} < \theta < \pi$. This means a *forward shift of the scattering* due to the *pressure of the incident light upon the outgoing photocurrent*, that was predicted by Wentzel [100].

The formula valid in each order was obtained by Fisher and Sauter (see [89, Vol. II]):

$$(14.19) \quad \mathbf{j}(t, \mathbf{x}) \sim \frac{\sin^2 \theta \cos^2 \varphi}{(1 - \beta \cos \theta)^4 |\mathbf{x}|^2} \mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \rightarrow \infty.$$

Remark 14.5 *It would be interesting to prove that the photoeffect is an inherent feature of the coupled Maxwell-Schrödinger equations and holds under the same spectral condition (14.1).*

14.6 Photoeffect for Excited States

The angular dependence (14.5) of the limiting amplitude is characteristic for the light scattering by spherically-symmetric ground state (12.1). The photoeffect is observed also in the light scattering by excited stationary states of atoms which are not spherically-symmetric. In this case, the formula of type (14.16) for the limiting amplitude also holds and admits a long-range asymptotics of type (14.2). However, its angular distribution is different from (14.5): see [89], where the distribution is obtained for the excited states from K -, L -, and M -shells of many-electron atoms.

15 Scattering of Particles: Rutherford Formula

Here, the scattering of charged particles at a heavy, charged object (“nucleus”) is considered. First, the classical case is considered, when the scattering particles are described by the trajectories of classical mechanics. Then, the quantum mechanical scattering of an electron with large momentum by a hydrogen atom is studied.

15.1 Classical Scattering by a Nucleus

Let us consider scattering of charged particles with charge Q and mass M by a nucleus with charge $|e|Z > 0$. We will derive the Rutherford formula in both the repulsive case, when $Q > 0$, and the attractive case, when $Q < 0$. The repulsive case corresponds, for example, to the scattering of α -particles with $Q = 2|e|$. The attractive case corresponds, for example, to the scattering of electrons with $Q = e < 0$. Let the heavy nucleus be situated at the origin. The incident particle moves along the trajectory $\mathbf{x} = \mathbf{x}(t)$, $-\infty < t < \infty$. We assume that the particles come from infinity, i.e.

$$(15.1) \quad \lim_{t \rightarrow -\infty} |\mathbf{x}(t)| = \infty.$$

Then the trajectory is a hyperbola as described, e.g., in Exercise 45.1. Let us choose the coordinates \mathbf{x}_k in space in such a way that $\mathbf{x}_3(t) \equiv 0$ and

$$(15.2) \quad \lim_{t \rightarrow -\infty} \dot{\mathbf{x}}(t) = (v, 0, 0), \quad \lim_{t \rightarrow -\infty} \mathbf{x}_1(t) = -\infty, \quad \lim_{t \rightarrow -\infty} \mathbf{x}_2(t) = b,$$

where b is the *impact* parameter. Let us further choose standard polar coordinates in the plane $\mathbf{x}_3 = 0$,

$$(15.3) \quad \mathbf{x}_1 = r \cos \theta, \quad \mathbf{x}_2 = r \sin \theta$$

and denote by $r(t), \theta(t)$ the trajectory of the particle in these coordinates. Then the initial scattering conditions (15.2) imply

$$(15.4) \quad \lim_{t \rightarrow -\infty} \theta(t) = \pi, \quad \lim_{t \rightarrow -\infty} r(t) \sin \theta(t) = b.$$

Angle of scattering

Let us calculate the final scattering angle

$$(15.5) \quad \bar{\theta} := \lim_{t \rightarrow -\infty} \theta(t).$$

Lemma 15.1 *The final scattering angle satisfies the equation*

$$(15.6) \quad \cot \frac{\bar{\theta}}{2} = \frac{Mbv^2}{Q|e|Z}.$$

Proof First, let us write the angular momentum and energy conservation:

$$(15.7) \quad r^2(t)\dot{\theta}(t) = bv, \quad \frac{M}{2}(\dot{r}^2(t) + r^2(t)\dot{\theta}^2(t)) + \frac{Q|e|Z}{r(t)} = \frac{M}{2}v^2.$$

Let us substitute here

$$(15.8) \quad \dot{r}(t) := \frac{dr}{dt} = \frac{dr}{d\theta} \frac{d\theta}{dt} = r'\dot{\theta}.$$

Then the energy conservation becomes,

$$(15.9) \quad \frac{M}{2}\dot{\theta}^2(t)(|r'(t)|^2 + r^2(t)) + \frac{Q|e|Z}{r(t)} = \frac{M}{2}v^2.$$

Now let us make the **Clerot substitution** $r = 1/\rho$. Then $r' = -\rho'/\rho^2$ and the momentum conservation gives $\dot{\theta}(t) = vb\rho^2$. Therefore, (15.9) reads,

$$(15.10) \quad \frac{M}{2}b^2v^2(|\rho'|^2 + \rho^2) + Q|e|Z\rho = \frac{M}{2}v^2.$$

Let us differentiate this expression in θ . Then after division by ρ' , we get the **Clerot equation**

$$(15.11) \quad \rho'' + \rho = C := -\frac{Q|e|Z}{Mb^2v^2}.$$

The general solution to this equation is

$$(15.12) \quad \rho(\theta) = A \cos \theta + B \sin \theta + C.$$

The initial scattering conditions (15.2) give,

$$(15.13) \quad \lim_{\theta \rightarrow \pi} \rho(\theta) = 0, \quad \lim_{\theta \rightarrow \pi} \frac{\rho(\theta)}{\sin(\theta)} = \frac{1}{b}.$$

Substituting here (15.12), we get $-A + C = 0$ and $B = 1/b$, hence

$$(15.14) \quad \rho(\theta) = C(1 + \cos \theta) + \frac{1}{b} \sin \theta.$$

At last, for the final scattering angle we get from (15.5) that $\rho(\bar{\theta}) = 0$. Hence,

$$(15.15) \quad C(1 + \cos \bar{\theta}) + \frac{1}{b} \sin \bar{\theta} = 0.$$

This implies (15.6). ■

Remark 15.2 *The solution $\bar{\theta} \in (0, \pi) \cup (\pi, 2\pi)$ to Equation (15.6) exists and is unique. The repulsive and attractive cases correspond to $Q > 0$, $\bar{\theta} \in (0, \pi)$ and $Q < 0$, $\bar{\theta} \in (\pi, 2\pi)$, respectively.*

Differential cross section

Now let us assume that the incident particles constitute a beam with a flux density of n particles per cm^2sec . Let us denote by $N = N(b, b + db)$ the number of incident particles per sec with an impact parameter within the interval $[b, b + db]$. By axial symmetry, we get for infinitesimal db ,

$$(15.16) \quad N(b, b + db) = n2\pi b db$$

The particles are scattered into the spatial angle $d\Omega = 2\pi \sin \bar{\theta} d\bar{\theta}$.

Definition 15.3 *The differential cross section of the scattering is defined by*

$$(15.17) \quad D(\bar{\theta}) := \frac{N/d\Omega}{n} = \frac{bdb}{\sin \bar{\theta} d\bar{\theta}}.$$

Let us calculate the cross section. Rewriting (15.6) in the form

$$(15.18) \quad b^2 = \left(\frac{Q|e|Z}{Mv^2} \right)^2 \cot^2 \frac{\bar{\theta}}{2}$$

and differentiating, we get

$$(15.19) \quad 2bdb = \left(\frac{Q|e|Z}{Mv^2} \right)^2 2 \cot \frac{\bar{\theta}}{2} \frac{1}{\sin^2 \frac{\bar{\theta}}{2}} \frac{d\bar{\theta}}{2}.$$

Substituting this into (15.17), we get the *Rutherford formula*

$$(15.20) \quad D(\bar{\theta}) = \frac{\left(\frac{Q|e|Z}{Mv^2} \right)^2}{4 \sin^4 \frac{\bar{\theta}}{2}}.$$

15.2 Quantum Scattering of Electrons by the Hydrogen Atom

We consider the scattering of the electron beam by a hydrogen atom in its ground state $\psi_1(t, \mathbf{x})$, (12.1). The incident electron beam is described by the plane wave (7.6):

$$(15.21) \quad \psi_{\text{in}}(t, \mathbf{x}) = C e^{i(\mathbf{k}\mathbf{x} - \omega t)}, \quad \mathbf{k} \neq 0,$$

which is a solution to the *free* Schrödinger equation, i.e., without an external Maxwell field. Hence, (see (7.7))

$$(15.22) \quad \hbar\omega = \frac{\hbar^2}{2\mu} \mathbf{k}^2 > 0.$$

Corresponding electric current density is given by (7.9):

$$(15.23) \quad \mathbf{j}_{\text{in}}(t, \mathbf{x}) := \frac{e}{\mu} [-i\hbar \nabla \psi_{\text{in}}(t, \mathbf{x})] \cdot \psi_{\text{in}}(t, \mathbf{x}) = \frac{e\hbar\mathbf{k}}{\mu} |C|^2.$$

We will assume that $|C| \ll C_1$ (see (12.1)) and consider the incident plane wave as a small perturbation. Hence, the total wave field, approximately, is a solution to the Schrödinger equation of the type (7.2),

$$(15.24) \quad [i\hbar\partial_t - e\phi(\mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]^2 \psi(t, \mathbf{x}),$$

where $\phi(\mathbf{x})$, $A(\mathbf{x})$ are the potentials of the total static Maxwell field corresponding to the ground state ψ_1 . In particular (cf. (12.3)),

$$(15.25) \quad -\frac{1}{4\pi} \Delta\phi(\mathbf{x}) = \rho(\mathbf{x}) = e|\psi_1(t, \mathbf{x})|^2 + |e|\delta(\mathbf{x}),$$

where $|e|\delta(\mathbf{x})$ is the charge density of the nucleus. The potential decays at infinity like $|x|^{-2}$, since $\int \rho(\mathbf{x})d\mathbf{x} = 0$. For simplicity of calculations, we assume first that

$$(15.26) \quad \phi(\mathbf{x}) = e \frac{e^{-\varepsilon|\mathbf{x}|}}{|\mathbf{x}|},$$

where $\varepsilon > 0$ is small. At the end we will perform the limit $\varepsilon \rightarrow 0$.

Radiated wave

Let us split the total wave field into three terms,

$$(15.27) \quad \psi(t, \mathbf{x}) = \psi_{\text{in}}(t, \mathbf{x}) + \psi_1(t, \mathbf{x}) + \psi_r(t, \mathbf{x}),$$

where $\psi_r(t, \mathbf{x})$ is a small radiated wave. Substituting (15.27) into the Schrödinger equation (7.2), we get

$$(15.28) \quad [i\hbar\partial_t - e\phi(\mathbf{x})][\psi_{\text{in}}(t, \mathbf{x}) + \psi_r(t, \mathbf{x})] = \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x})]^2[\psi_{\text{in}}(t, \mathbf{x}) + \psi_r(t, \mathbf{x})],$$

since $\psi_1(t, \mathbf{x})$ is an exact solution. Neglecting the "relativistic" corrections due to the small term $\frac{e}{c}\mathbf{A}(\mathbf{x})$, we rewrite the equation as

$$(15.29) \quad \left(i\hbar\partial_t - e\phi(\mathbf{x}) - \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}}]^2\right)\psi_r(t, \mathbf{x}) = e\phi(\mathbf{x})\psi_{\text{in}}(t, \mathbf{x}) = e\phi(\mathbf{x})Ce^{i(\mathbf{k}\mathbf{x} - \omega t)}.$$

Since $\omega > 0$ by (15.22), the frequency ω belongs to the continuous spectrum of the Schrödinger operator. Now let us apply the limiting amplitude principle (43.28) to derive the long-time asymptotics of the solution $\psi_r(t, \mathbf{x})$. Up to a contribution of the discrete spectrum (see Remark 43.11), the solution admits the asymptotics

$$(15.30) \quad \psi_r(t, \mathbf{x}) \sim \psi_r(\mathbf{x})e^{-i\omega t}, \quad t \rightarrow \infty.$$

Substituting this asymptotics into Equation (15.29), we get a stationary equation for the limiting amplitude

$$(15.31) \quad \left(\hbar\omega - e\phi(\mathbf{x}) - \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}}]^2\right)\psi_r(\mathbf{x}) = eC\phi(\mathbf{x})e^{i\mathbf{k}\mathbf{x}}.$$

Neglecting the term with the spatial decay at the LHS, we finally get the Helmholtz equation

$$(15.32) \quad \left(\mathbf{k}^2 + \nabla_{\mathbf{x}}^2\right)\psi_r(\mathbf{x}) = \frac{2\mu eC}{\hbar^2}\phi(\mathbf{x})e^{i\mathbf{k}\mathbf{x}},$$

since $\mathbf{k}^2 = \frac{2\mu\omega}{\hbar} > 0$ by (15.22). This is an equation of the type (14.11). Therefore, the solution is given by a convolution similar to (14.12),

$$(15.33) \quad \psi_r(\mathbf{x}) = -\frac{2\mu eC}{\hbar^2} \int \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} \phi(\mathbf{y})e^{iky} d\mathbf{y}, \quad k := |\mathbf{k}|.$$

The convolution (15.33) is almost identical to the last integral (12.12) if we identify $\mathbf{k} = (0, 0, k)$. Evaluating this by the method (12.13)-(12.16), we get

$$(15.34) \quad \psi_r(\mathbf{x}) \sim -C \frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|} f(k, \theta), \quad |\mathbf{x}| \rightarrow \infty,$$

where

$$(15.35) \quad f(k, \theta) = \frac{2\mu e}{\hbar^2} \int_0^\infty \frac{\sin K|\mathbf{y}|}{K|\mathbf{y}|} \phi(\mathbf{y})|\mathbf{y}|^2 d|\mathbf{y}|, \quad K := 2k \sin \frac{\theta}{2}.$$

Now (15.30) becomes for large $|\mathbf{x}|$

$$(15.36) \quad \psi_r(t, \mathbf{x}) \sim -C \frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|} f(k, \theta) e^{-i\omega t}, \quad t \rightarrow \infty.$$

The limiting amplitude ψ_r has a slow decay at infinity and infinite energy. This corresponds to the fact that the frequency $\omega > 0$ belongs to the continuous spectrum (see Remark 43.10 ii). Physically this describes the radiation of electrons to infinity as we will see below.

Differential cross section

Let us calculate the electric current corresponding to the radiated wave $\psi_r(t, \mathbf{x})$. As in (15.23), we have for large $|\mathbf{x}|$,

$$(15.37) \quad \mathbf{j}_r(t, \mathbf{x}) := \frac{e}{\mu} [-i\hbar \nabla \psi_r(t, \mathbf{x})] \cdot \psi_r(t, \mathbf{x}) \sim \frac{e\hbar k \mathbf{n}(\mathbf{x})}{\mu} \frac{|f(k, \theta)|^2}{|\mathbf{x}|^2} |C|^2, \quad \mathbf{n}(\mathbf{x}) := \frac{\mathbf{x}}{|\mathbf{x}|}.$$

Let us note that the current at infinity is radial.

Definition 15.4 *The differential cross section of the scattering is defined by*

$$(15.38) \quad D(\mathbf{n}) := \lim_{\mathbf{x}/|\mathbf{x}|=\mathbf{n}, |\mathbf{x}| \rightarrow \infty} \frac{|\mathbf{j}_r(t, \mathbf{x})|}{|\mathbf{j}_{in}(t, \mathbf{x})|} |\mathbf{x}|^2, \quad \mathbf{n} \in \mathbb{R}^3.$$

Remark 15.5 *The definition obviously agrees with Definition 15.3.*

Now (15.37) and (15.23) imply that

$$(15.39) \quad D(\mathbf{n}) = |f(k, \theta)|^2.$$

Finally, substituting (15.26) into (15.35), we get

$$(15.40) \quad f(k, \theta) = \frac{2\mu e^2}{K\hbar^2} \int_0^\infty \sin K|\mathbf{y}| e^{-\varepsilon|\mathbf{y}|} d|\mathbf{y}| = \frac{2\mu e^2}{\hbar^2(K^2 + \varepsilon^2)}.$$

Here $K = 2k \sin \frac{\theta}{2} \gg \varepsilon^2$ for $\theta \neq 0$ and large values of $k > k(\theta)$. Then we can drop ε^2 in the limit $\varepsilon \rightarrow 0$ and obtain

$$(15.41) \quad f(k, \theta) = \frac{2\mu e^2}{\hbar^2 K^2} = \frac{\mu e^2}{2\hbar^2 k^2 \sin^2 \frac{\theta}{2}}.$$

We rewrite this expression using (7.10):

$$(15.42) \quad f(k, \theta) = \frac{e^2}{2\mu \mathbf{v}^2 \sin^2 \frac{\theta}{2}}.$$

Now (15.39) implies the formula

$$(15.43) \quad D(\mathbf{n}) = \frac{\left(\frac{e^2}{\mu \mathbf{v}^2}\right)^2}{4 \sin^4 \frac{\theta}{2}},$$

which coincides with the classical Rutherford formula (15.20) with $Q = e$ and $Z = 1$.

Remark 15.6 *Born considered the agreement of (15.43) with the classical formula (15.20) as the crucial confirmation of the probabilistic interpretation of the wave function: $|\psi(t, \mathbf{x})|$ is the **density of the probability of the particle registration**, and the expression (15.37) is e times the **flux of the probability**, [12].*

16 Hydrogen Atom in a Magnetic Field. Normal Zeemann Effect

We derive quantum stationary states and the corresponding energies for a hydrogen atom in a uniform magnetic field. We also analyze the normal Zeemann effect.

Let us consider a hydrogen atom in an external static magnetic field $\mathbf{B}(\mathbf{x})$ with vector potential $\mathbf{A}(\mathbf{x})$. Then the Lorentz gauge condition (4.5) is equivalent to the Coulomb gauge condition

$$(16.1) \quad \nabla_{\mathbf{x}} \mathbf{A}(\mathbf{x}) = 0.$$

The Schrödinger equation (6.2) becomes

$$(16.2) \quad i\hbar \partial_t \psi(t, \mathbf{x}) = \frac{1}{2\mu} (-i\hbar \nabla_{\mathbf{x}} - \frac{e}{c} \mathbf{A}(\mathbf{x}))^2 \psi(t, \mathbf{x}) - \frac{e^2}{|\mathbf{x}|} \psi(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Evaluating, we get by (16.1)

$$(16.3) \quad \begin{aligned} i\hbar \partial_t \psi(t, \mathbf{x}) = & -\frac{1}{2\mu} \hbar^2 \Delta \psi(t, \mathbf{x}) + i \frac{\hbar e}{\mu c} \mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \psi(t, \mathbf{x}) \\ & - \frac{1}{2\mu} \hbar^2 \mathbf{A}^2(\mathbf{x}) \psi(t, \mathbf{x}) - \frac{e^2}{|\mathbf{x}|} \psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3. \end{aligned}$$

Let us assume that the potential $\mathbf{A}(\mathbf{x})$ is small:

$$(16.4) \quad |\mathbf{A}(\mathbf{x})| \ll 1.$$

Then we can neglect the term with $\mathbf{A}^2(\mathbf{x})$ in (16.3) and get the equation

$$(16.5) \quad i\hbar \partial_t \psi(t, \mathbf{x}) = -\frac{1}{2\mu} \hbar^2 \Delta \psi(t, \mathbf{x}) + i \frac{\hbar e}{\mu c} \mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \psi(t, \mathbf{x}) - \frac{e^2}{|\mathbf{x}|} \psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

16.1 Uniform Magnetic Field

Now let us consider a hydrogen atom in an external uniform static magnetic field \mathbf{B} with vector potential $\mathbf{A}(\mathbf{x}) = \mathbf{B} \times \mathbf{x} / 2$. Let us choose the coordinates in such a way that $\mathbf{B} = (0, 0, B)$ with $B \geq 0$. Then the vector potential is given by $\mathbf{A}(\mathbf{x}) = \frac{1}{2} B (-x_2, x_1, 0) = \frac{1}{2} B |\mathbf{x}| \sin \theta \mathbf{e}_\varphi$. Therefore, we have the *static* Maxwell field with the potentials

$$(16.6) \quad \left. \begin{aligned} \phi(\mathbf{x}) &:= -\frac{e}{|\mathbf{x}|}, \\ \mathbf{A}(\mathbf{x}) &:= \frac{1}{2} B (-x_2, x_1, 0) \end{aligned} \right| \mathbf{x} \in \mathbb{R}^3.$$

Remark 16.1 *The angular momentum conservation (6.13) holds for $n = 3$:*

$$(16.7) \quad \mathbf{L}_3(t) := \langle \psi(t, \mathbf{x}), \hat{\mathbf{L}}_3 \psi(t, \mathbf{x}) \rangle = \text{const},$$

where $\hat{\mathbf{L}}_3$ is the operator in (6.18) with $n = 3$. The conservation reflects the axial symmetry of Equation (16.5) with the potentials (16.6).

Let us note that $i\hbar\mathbf{A}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} = B i\hbar \nabla_{\varphi} / 2 = -B \hat{\mathbf{L}}_3 / 2$ by (6.18). Here, φ is the angular coordinate of the rotation around the vector $\mathbf{e}_3 := (0, 0, 1)$, i.e. around \mathbf{B} . Then (16.5) becomes

$$(16.8) \quad i\hbar \partial_t \psi(t, \mathbf{x}) = -\frac{1}{2\mu} \hbar^2 \Delta \psi(t, \mathbf{x}) - \frac{e}{2\mu c} B \hat{\mathbf{L}}_3 \psi(t, \mathbf{x}) - \frac{e^2}{|\mathbf{x}|} \psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

Hence, the corresponding stationary equation reads

$$(16.9) \quad E_{\omega} \psi_{\omega}(\mathbf{x}) = -\frac{1}{2\mu} \hbar^2 \Delta \psi_{\omega}(\mathbf{x}) - \omega_{\mathcal{L}} \hat{\mathbf{L}}_3 \psi_{\omega}(\mathbf{x}) - \frac{e^2}{|\mathbf{x}|} \psi_{\omega}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3,$$

where $\omega_{\mathcal{L}} := eB/(2\mu c) < 0$ is the **Larmor frequency**. Therefore, Theorem 8.1, Theorem 8.4 ii) and (8.32) imply the following theorem.

Theorem 16.2 *i) The quantum stationary states ψ_{ω} of the hydrogen atom in a uniform magnetic field coincide with the functions ψ_{lmn} from (8.34).*

ii) The corresponding energies are equal to

$$(16.10) \quad E_{mn} := -2\pi\hbar R/n^2 + m\hbar\omega_{\mathcal{L}}, \quad n = 1, 2, 3, \dots, \quad m = -l, \dots, l, \quad l \leq n - 1.$$

Radial scalar potential Let us consider more general equations of the type (16.6) with a static radial external potential $\phi(|x|)$ and the static uniform magnetic field B . Then the corresponding eigenvalue problem of the type (16.8) becomes

$$(16.11) \quad E_{\omega} \psi_{\omega}(\mathbf{x}) = -\frac{1}{2\mu} \hbar^2 \Delta \psi_{\omega}(\mathbf{x}) - i\hbar\omega_{\mathcal{L}} \nabla_{\varphi} \psi_{\omega}(\mathbf{x}) + e\phi(|\mathbf{x}|) \psi_{\omega}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

The following theorem can be proved by the same arguments as in Theorems 8.1, 16.2.

Theorem 16.3 *i) The quantum stationary states of the electron in a static central electric and a uniform magnetic field have the following form in spherical coordinates (cf. (8.34)):*

$$(16.12) \quad \psi_{lmn}^* = R_{ln}(r) F_l^m(\theta) e^{im\varphi}.$$

ii) The corresponding energies are equal to

$$(16.13) \quad E_{lmn}^* = E_{ln}^* + m\hbar\omega_{\mathcal{L}}, \quad n = 1, 2, 3, \dots, \quad m = -l, \dots, l, \quad l \leq n - 1,$$

where E_{ln}^* are the energies corresponding to $B = 0$.

16.2 Normal Zeemann Effect: Triplet Spectrum

In 1895 Zeemann observed the influence of a magnetic field on the spectral lines of atoms and molecules. Consider the spectrum of radiation of the hydrogen atom in a uniform magnetic field. Let us calculate the level splitting and the polarization of the radiation emitted by the atom: we will see that both are as in the classical description (see Exercise 45.9). Namely, the *unperturbed spectral lines*, corresponding to the zero magnetic field $B = 0$, split into the *normal triplet* in the case $B \neq 0$.

First, by Theorem 16.2, the bound state eigenfunctions are (8.34):

$$(16.14) \quad \psi_k = R_{nl}(r) F_l^m(\theta) e^{im\varphi}, \quad k = (nlm),$$

and the eigenvalues are

$$(16.15) \quad \omega_k = -\frac{2\pi R}{n^2} + m\omega_{\mathcal{L}}, \quad \omega_{\mathcal{L}} = -\frac{eB}{2\mu c}.$$

In the case $B = 0$, we have also $\omega_{\mathcal{L}} = 0$. Hence, the *unperturbed spectral lines*, $\omega_{kk'}^0 = 2\pi R[\frac{1}{n'^2} - \frac{1}{n^2}]$, where $k = (nlm)$ and $k' = (n'l'm')$.

Next consider the case $B \neq 0$. The calculations (10.18)-(10.22) imply the formula for the intensity of the spectral line $\omega_{kk'}$ in the dipole approximation

$$(16.16) \quad \mathbf{J}_{kk'} \sim \int_0^\infty R_{nl}(r)\overline{R_{n'l'}}(r)r^3 dr \int_S \mathbf{x}F_l^m(\theta)\overline{F_{l'}^{m'}}(\theta)e^{i(m-m')\varphi} dS.$$

We know this integral is nonzero only for $m' = m, m \pm 1$ and $l' = l \pm 1$. Hence, for $B \neq 0$, the unperturbed spectral line $\omega_{kk'}^0$ generates the *triplet* of spectral lines with frequencies

$$(16.17) \quad \begin{cases} m' = m & : \omega_{kk'} = \omega_{kk'}^0 \\ m' = m - 1 & : \omega_{kk'} = \omega_{kk'}^0 + \omega_{\mathcal{L}} \\ m' = m + 1 & : \omega_{kk'} = \omega_{kk'}^0 - \omega_{\mathcal{L}}, \end{cases}$$

precisely like in the classical case. To find the polarizations we re-write the vector \mathbf{x} like

$$(16.18) \quad \mathbf{x} = \sin\theta(e^{-i\varphi}\mathbf{e}_+ + e^{i\varphi}\mathbf{e}_-) + \cos\theta\mathbf{e}_z, \quad \mathbf{e}_\pm = \frac{1}{2}(\hat{e}_x \pm i\hat{e}_y).$$

Inserting back into (16.16) we find that for $m' = m$ only the \mathbf{e}_z component contributes, and for $m' = m \pm 1$ the \mathbf{e}_\mp component. Therefore, we find from (10.12) the radiation in the form

$$(16.19) \quad \begin{cases} m' = m & : \mathbf{A}_{kk'}(t, \mathbf{x}) \sim \frac{1}{|\mathbf{x}|} \text{Re} c_k \bar{c}_{k'} \mathbf{e}_z e^{-i\omega_0(t-|\mathbf{x}|/c)} \\ m' = m - 1 & : \mathbf{A}_{kk'}(t, \mathbf{x}) \sim \frac{1}{|\mathbf{x}|} \text{Re} c_k \bar{c}_{k'} \mathbf{e}_+ e^{-i(\omega_0 + \omega_{\mathcal{L}})(t-|\mathbf{x}|/c)} \\ m' = m + 1 & : \mathbf{A}_{kk'}(t, \mathbf{x}) \sim \frac{1}{|\mathbf{x}|} \text{Re} c_k \bar{c}_{k'} \mathbf{e}_- e^{-i(\omega_0 - \omega_{\mathcal{L}})(t-|\mathbf{x}|/c)}. \end{cases}$$

Note that the resulting polarizations of the radiation fields are exactly like in the classical case (see Exercise 45.9).

Remark 16.4 The spectral lines (16.17) do not depend on the *azimuthal quantum number* l as well as the eigenvalues (16.15). Therefore, the radiations, induced by the pairs $k = (mln), k' = (m'l'n')$ with $m' = m, m \pm 1$ and all possible azimuthal quantum number l, l' , contribute to the same frequencies (16.17). On the other hand, in many cases the atom spectra in a magnetic field demonstrate the *multiplet structure* drastically different from the triplet structure (16.17). This **anomalous Zeemann effect** cannot be explained by the Schrödinger equation (16.2). The explanation is provided by the **Pauli** equation which takes into account the **electron spin** (see Section 22.3).

17 Diamagnetism and Paramagnetism

We compute a magnetic moment of a hydrogen atom and split it in two parts, one of which describes a diamagnetism of the atom and the other describes a paramagnetism. This allows us to calculate the *magnetic susceptibility* of the hydrogen atom (see Lecture 14).

Namely, we will consider the stationary states of the atom in the uniform external magnetic field \mathbf{B} and calculate their magnetic moment (42.22)

$$(17.1) \quad \mathbf{m} = \frac{1}{2c} \int \mathbf{y} \times \mathbf{j}(\mathbf{y}) d\mathbf{y}.$$

Let us choose the coordinates in such a way that $\mathbf{B} = (0, 0, B)$ with $B \geq 0$. Then the corresponding vector potential is given by $\mathbf{A}(\mathbf{x}) = \frac{1}{2}B(-x_2, x_1, 0) = \frac{1}{2}B|\mathbf{x}| \sin \theta \mathbf{e}_\varphi$.

17.1 Electric Current at Stationary States

First we calculate the electric current \mathbf{j} defined by the Born approximation (7.28),

$$(17.2) \quad \mathbf{j}(t, \mathbf{x}) = \frac{e}{\mu} [-i\hbar \nabla - \frac{e}{c} \mathbf{A}(\mathbf{x})] \psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}).$$

The stationary states are given by (8.34) (see Theorem 16.2). Let us write (8.34) in the form

$$(17.3) \quad \psi(t, \mathbf{x}) = a_{lmn}(r, \theta) e^{im\varphi} e^{-i\omega t},$$

where a_{lmn} is a real function according to Theorem 8.4 ii). Let us express the gradient operator in spherical coordinates (see (9.15)):

$$(17.4) \quad \nabla = \mathbf{e}_r \nabla_r \psi + \mathbf{e}_\theta \frac{\nabla_\theta}{r} + \mathbf{e}_\varphi \frac{\nabla_\varphi}{r \sin \theta}.$$

Then (17.2) and (17.3) give,

$$(17.5) \quad \begin{aligned} \mathbf{j}(t, \mathbf{x}) = & \frac{e}{\mu} [-i\hbar \mathbf{e}_r \nabla_r a_{lmn}(r, \theta)] \cdot a_{lmn}(r, \theta) + \frac{e}{\mu} [-i\hbar \mathbf{e}_\theta \frac{\nabla_\theta}{r} a_{lmn}(r, \theta)] \cdot a_{lmn}(r, \theta) \\ & + \frac{e}{\mu} [m\hbar \mathbf{e}_\varphi \frac{1}{r \sin \theta}] \psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}) - \frac{e^2}{\mu c} \mathbf{A}(\mathbf{x}) \psi(t, \mathbf{x}) \cdot \psi(t, \mathbf{x}). \end{aligned}$$

The first and second term on the RHS are zero since the function $a_{lmn}(r, \theta)$ is real. Hence finally,

$$(17.6) \quad \mathbf{j}(t, \mathbf{x}) = \frac{e}{\mu} [m\hbar \mathbf{e}_\varphi \frac{1}{r \sin \theta}] |\psi(t, \mathbf{x})|^2 - \frac{e^2}{\mu c} B |\mathbf{x}| \sin \theta \mathbf{e}_\varphi |\psi(t, \mathbf{x})|^2 = \mathbf{j}'(\mathbf{x}) + \mathbf{j}''(\mathbf{x}).$$

The magnetic moment (17.1) becomes

$$(17.7) \quad \mathbf{m} = \frac{1}{2c} \int \mathbf{y} \times \mathbf{j}'(\mathbf{y}) d\mathbf{y} + \frac{1}{2c} \int \mathbf{y} \times \mathbf{j}''(\mathbf{y}) d\mathbf{y} = \mathbf{m}' + \mathbf{m}''.$$

The currents $\mathbf{j}'(\mathbf{x})$, $\mathbf{j}''(\mathbf{x})$ are axially symmetric with respect to the axis Ox_3 , and

$$(17.8) \quad (\mathbf{y} \times \mathbf{e}_\varphi)_3 = |\mathbf{y}| \mathbf{e}_3 \sin \theta.$$

17.2 Langevin Formula for Diamagnetic Susceptibility

By (17.8), the direction of \mathbf{m}'' is opposite to \mathbf{B} , hence, \mathbf{m}'' describes the *diamagnetism* of the atom. The corresponding diamagnetic susceptibility (per atom) is defined by $\mathbf{m}'' = \chi_m \mathbf{B}$, hence

$$(17.9) \quad \chi_m := -\frac{|\mathbf{m}''|}{B} = -\frac{e^2}{2\mu c^2} \int |\mathbf{y}|^2 \sin^2 \theta |\psi|^2 d\mathbf{y} = -\frac{e^2}{2\mu c^2} \Theta_3,$$

where Θ_3 denotes the moment of inertia of the distribution $|\psi|^2$ with respect to the axis $O\mathbf{x}_3$. For spherically symmetric stationary states we have

$$(17.10) \quad \Theta_3 = \frac{1}{3} \bar{\Theta}, \quad \bar{\Theta} := \int |\mathbf{y}|^2 |\psi|^2 d\mathbf{y}.$$

Then (17.9) becomes the *Langevin formula* (cf. Exercise 45.10)

$$(17.11) \quad \chi_m = -\frac{e^2 \bar{\Theta}}{6\mu c^2}.$$

For spherically non-symmetric states the formula also holds *in the mean* due to the *random orientation* of the atoms with respect to the direction of the magnetic field.

17.3 Paramagnetism

The moment \mathbf{m}' is different from zero even in the absence of the magnetic field,

$$(17.12) \quad \mathbf{m}' = m\hbar \frac{e}{2c\mu} \int \frac{(\mathbf{y} \times \mathbf{e}_\varphi)_3 \mathbf{e}_3}{|\mathbf{y}| \sin \theta} |\psi(t, \mathbf{y})|^2 d\mathbf{y} = m\hbar \mathbf{e}_3 \frac{e}{2c\mu} \int |\psi(t, \mathbf{y})|^2 d\mathbf{y} = \frac{e}{2c\mu} m\hbar \mathbf{e}_3$$

by (17.8). Hence, \mathbf{m}' describes the *paramagnetism* of the atom. However, its mathematical expectation is zero due to the random orientation of the atom since the probabilities of the values m and $-m$ are identical. On the other hand, the probabilities are not identical if the magnetic field does not vanish. The corresponding *statistical* theory has been developed by Langevin.

Part IV

Electron Spin and Pauli Equation

18 Electron Spin: Experiments and Interpretation

We describe some experiments which have inspired the ideas of a *spin* and a *magnetic moment* of an electron, namely, the Einstein-de-Haas experiment, the Stern-Gerlach experiment, and the anomalous Zeemann effect.

The concept of spin and magnetic moment of the electron has been introduced by Goudsmith and Uhlenbeck in 1925: *Every electron has an intrinsic angular momentum (spin) \mathbf{s} with the magnitude $|\mathbf{s}| = \frac{\hbar}{2}$, and a magnetic moment \mathbf{m}_s with the magnitude $|\mathbf{m}_s| = \mu_B := \frac{|e|\hbar}{2\mu c}$, which is the Bohr magneton.* Here the term “intrinsic” means that the spin angular momentum is not related to a rotation of the particle, and spin magnetic moment is not related to the corresponding convection current.

Let us consider the experimental data which inspired this conjecture.

18.1 Einstein-de Haas Experiment

Let us consider an iron bar positioned vertically in the earth’s gravitational field, which is attached to a vertical string in such a way that it can rotate about its axis. Let us magnetize the bar by a vertical external weak magnetic field. The field orients the elementary Ampere molecular currents, so that their magnetic moments increase the external magnetic field. Therefore, the corresponding elementary angular momenta are also oriented in the same direction, and the sum of the microscopic angular momenta increases. By axial symmetry, the total angular momentum of the bar, macroscopic+microscopic, is conserved. Hence, the bar as a whole must change its macroscopic angular momentum, too. In 1915 Einstein and de Haas measured the changes of magnetic moment and macroscopic angular momentum of the bar to check their ratio. The classical and quantum theories predict the ratio $\frac{e}{2\mu c}$. However, the experimental results contradict this value. This suggests the existence of an additional magnetic moment of the atoms which is responsible for the anomalous ratio.

Classical theory

Let us assume that in each molecule, the currents are caused by the rotation of electrons with the same angular velocity ω . Then the angular momentum and magnetic moment of the molecule are given by

$$(18.1) \quad \mathbf{L} := \sum \mathbf{x}_k \times \mu(\omega \times \mathbf{x}_k), \quad \mathbf{m} := \frac{1}{2c} \sum \mathbf{x}_k \times e(\omega \times \mathbf{x}_k),$$

where \mathbf{x}_k are the electron positions. Therefore, the following key relation holds for each molecule,

$$(18.2) \quad \mathbf{m} = \frac{e}{2\mu c} \mathbf{L}.$$

Hence, an increment of the magnetization, $\Delta \mathbf{m}_3$, of the bar, is followed by the corresponding increment, $-\Delta \mathbf{L}_3 = -\frac{e}{2\mu c} \Delta \mathbf{m}_3$, of its macroscopic angular momentum. Both quantities, $\Delta \mathbf{m}_3$ and $\Delta \mathbf{L}_3$, can be measured experimentally: $-\Delta \mathbf{L}_3$ by the torsion vibration of the string, and $\Delta \mathbf{m}_3$ by the residual magnetism. The experiment was performed by Einstein and de Haas in 1915. However, the result was

$$(18.3) \quad \mathbf{m}_3 = g \frac{e}{2\mu c} \mathbf{L}_3$$

with the *Landé factor* $g \approx 2$, which contradicts (18.2). This contradiction inspired the ratio $|\mathbf{s}|/|\mathbf{m}_s| = \frac{e}{\mu c}$ in the Stern-Gerlach conjecture. This ratio corresponds to the Landé factor $g = 2$ and allows to explain the result of the Einstein-de Haas experiment.

Quantum theory

Let us check that the relation (18.2) also holds for the Schrödinger equation (16.8) for small B . Indeed, (17.1) and (17.2) imply for small B ,

$$\begin{aligned} \mathbf{m} &\approx \frac{1}{2c} \int \mathbf{y} \times \frac{e}{\mu} [-i\hbar \nabla_{\mathbf{y}}] \psi(t, \mathbf{y}) \cdot \psi(t, \mathbf{y}) d\mathbf{y} \\ (18.4) \qquad &= \frac{e}{2\mu c} \int [-i\hbar \mathbf{y} \times \nabla_{\mathbf{y}}] \psi(t, \mathbf{y}) \cdot \psi(t, \mathbf{y}) d\mathbf{y} = \frac{e}{2\mu c} \mathbf{L}. \end{aligned}$$

In particular, for the stationary state (17.3) we have the magnetization (cf. (17.12))

$$(18.5) \qquad \mathbf{m}_3 \approx \frac{e}{2\mu c} m \hbar.$$

On the other hand, the energy for the stationary state is given by (16.13). The external magnetic field causes the transitions to stationary states with lower values of the energy (16.13), which corresponds to greater values of m since $\omega_{\mathcal{L}} < 0$. Therefore, the magnetization (17.12) increases in the transitions.

18.2 Double Splitting

An additional suggestion for the spin is provided by the observation of the splitting of stationary states of atoms and molecules in a magnetic field B .

Stern-Gerlach experiment

In 1922 Stern and Gerlach sent a beam of silver atoms through an inhomogeneous magnetic field. Later similar experiments have been performed with hydrogen atoms. The atoms are in the ground state (12.1), which is non-degenerate according to the Schrödinger equation (16.8). Let us write (16.8) in the form

$$(18.6) \qquad i\hbar \partial_t \psi(t, \mathbf{x}) = -\frac{1}{2\mu} \hbar^2 \Delta \psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|) \psi(t, \mathbf{x}) - \frac{e}{2\mu c} \hat{\mathbf{L}} \mathbf{B} \psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

The interaction term $\hat{\mathbf{L}} \mathbf{B} \psi(t, \mathbf{x})$ vanishes for the ground state. Therefore, the ground state also satisfies Equation (18.6) with $\mathbf{B} \neq 0$. However, Stern and Gerlach observed a splitting of the beam into two components. This obviously contradicts the identity of all atoms in the ground state and suggests that

- i) For $\mathbf{B} = 0$ the eigenspace corresponding to the ground state has a dimension at least two.
- ii) For $\mathbf{B} \neq 0$ the eigenspace splits in two distinct eigenspaces.

This suggests the existence of an additional magnetic moment of the atoms which is responsible for the splitting in the magnetic field.

Anomalous Zeemann effect

The Schrödinger equation gives a satisfactory explanation of the 'normal' Zeemann effect with splitting into three lines. However, Zeemann demonstrated in 1895 that, for $B \neq 0$, most of the spectral lines are split into a different number of lines: two, five, etc. This *anomalous* Zeemann effect contradicts the Schrödinger theory, which only predicts the normal Zeemann effect. This also suggests the existence of an additional magnetic moment responsible for the anomalous splitting.

19 Pauli Equation

The Einstein-de Haas, Stern-Gerlach and anomalous Zeemann effects demonstrate that the Schrödinger equation requires a modification. To the Schrödinger equation, we introduce an additional spin magnetic moment corresponding to spin angular momentum. Then the Schrödinger equation becomes the Pauli equation which explains the Stern-Gerlach effect. Let us analyze the details.

19.1 Additional Magnetic Moment

Relation (18.4) implies that, for small $|\mathbf{B}|$, the last term in (18.6) can be rewritten as

$$(19.1) \quad -\frac{e}{2\mu c} \hat{\mathbf{L}}\mathbf{B}\psi(t, \mathbf{x}) \approx -\hat{\mathbf{m}}\mathbf{B}\psi(t, \mathbf{x}),$$

where $\hat{\mathbf{m}}$ is the quantum observable corresponding to the magnetic moment. This corresponds to the fact that $-\mathbf{m}\mathbf{B}$ is the energy of the magnetic moment \mathbf{m} in the magnetic field \mathbf{B} . The Stern-Gerlach experiment demonstrates that the magnetic moment \mathbf{m} does not vanish even for the ground state of the hydrogen atom. Therefore, we have to postulate the existence of an additional magnetic moment $\mathbf{m}_s \neq 0$. The Einstein-de Haas experiment suggests the Goudsmith-Uhlenbeck conjecture that the additional magnetic moment corresponds to an additional *spin angular momentum* of the electron, $\mathbf{s} = \text{spin}$, with the Landé factor 2, i.e.

$$(19.2) \quad \mathbf{m}_s = 2\frac{e}{2\mu c}\mathbf{s}.$$

Next, we have to specify a mathematical construction for this spin angular momentum \mathbf{s} .

19.2 Additional Angular Momentum

Orbital angular momentum

First, let us analyze the *orbital angular momentum* $\mathbf{L} = (\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3)$ for the case $A_k = 0$, $k = 1, 2, 3$ and spherically symmetric scalar potential $\phi(|\mathbf{x}|)$:

$$(19.3) \quad i\hbar\partial_t\psi(t, \mathbf{x}) = \mathcal{H}\psi(t, \mathbf{x}) := -\frac{1}{2\mu}\hbar^2\Delta\psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|)\psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

The angular momentum is a conserved quantity which corresponds to an invariance of the Lagrangian (2.6) with respect to the **regular representation** of the rotation group $SO(3)$. The regular representation R_g acts on the phase space $\mathcal{E} := L^2(\mathbb{R}^3)$ by the formula $R_g\psi(\mathbf{x}) := \psi(g\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$, for $g \in SO(3)$, $\psi \in \mathcal{E}$. The Schrödinger operator \mathcal{H} commutes with the representation and with the generators \mathbf{H}_k of the rotations around the axis $O\mathbf{x}_k$ (see (8.5)). The generators automatically satisfy the commutation relations (9.2) of the Lie algebra of the rotation group $SO(3)$. The corresponding conserved quantities are given by the Noether theorem and have the form (6.13):

$$(19.4) \quad \mathbf{L}_k = \langle \psi, \hat{\mathbf{L}}_k \psi \rangle, \quad k = 1, 2, 3,$$

where $\hat{\mathbf{L}}_k = \hbar\mathbf{H}_k$. An alternative proof of the conservation follows from the commutation $[\mathcal{H}, \mathbf{H}_k] = 0$ (see (9.2)).

Spin angular momentum

The analysis suggests that the spin angular momentum corresponds in a similar way to another action of the rotation group, which is different from the regular representation. To construct this **spinor representation**, we denote its generators by \mathbf{h}_k . Then similarly to (9.2), we necessarily have

$$(19.5) \quad [\mathbf{h}_1, \mathbf{h}_2] = i\mathbf{h}_3, \quad [\mathbf{h}_2, \mathbf{h}_3] = i\mathbf{h}_1, \quad [\mathbf{h}_3, \mathbf{h}_1] = i\mathbf{h}_2.$$

Further, define $\hat{\mathbf{s}}_k = -\hbar\mathbf{h}_k$ and the spin angular momenta

$$(19.6) \quad \mathbf{s}_k = \langle \psi, \hat{\mathbf{s}}_k \psi \rangle, \quad k = 1, 2, 3.$$

The momenta are conserved quantities if the modified Schrödinger operator commutes with \mathbf{h}_k .

The double splitting of the ground state in the Stern-Gerlach experiment with the hydrogen atom suggests that the dimension of the unperturbed ‘ground eigenspace’ is two. Therefore, in analogy with the orbital angular momentum, the ground state **formally** corresponds to $2l + 1 = 2$, i.e. $l = 1/2$, and, for $\mathbf{B} \neq 0$, it generates two stationary states which are eigenfunctions of the operator \mathbf{h}_3 with the eigenvalues $\pm 1/2$. Therefore, the action of the group $SO(3)$ in the unperturbed ground eigenspace is an irreducible two-dimensional representation S_g .

Such a two-dimensional representation is given by Proposition 9.8: it corresponds to the generators $\mathbf{h}_k := \frac{1}{2}\sigma_k$, where σ_k are the **Pauli matrices** (see (9.11))

$$(19.7) \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Exercise 19.1 Check the commutation relations (19.5) for the generators $\mathbf{h}_k = \frac{1}{2}\sigma_k$.

The simplest way to make the ground eigenspace two-dimensional is to define the modified phase space as the tensor product $\mathcal{E}^\otimes := \mathcal{E} \otimes \mathbb{C}^2$ and consider the tensor product of the regular representation T_g in \mathcal{E} and the spinor representation S_g in \mathbb{C}^2 .

Remark 19.2 By the definition of the tensor product of representations, all generators \mathbf{H}_k commute with all generators \mathbf{h}_j .

19.3 Pauli Equation. Uniform Magnetic Field

Let us summarize our discussion and change the Schrödinger equation (18.6) to the **Pauli equation** (1.14),

$$(19.8) \quad \begin{aligned} i\hbar\partial_t\Psi(t, \mathbf{x}) &= \mathcal{P}\Psi(t, \mathbf{x}) \\ &:= -\frac{1}{2\mu}\hbar^2\Delta\Psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|)\Psi(t, \mathbf{x}) - \frac{e}{2\mu c}\hat{\mathbf{L}}\mathbf{B}\Psi(t, \mathbf{x}) - \frac{e}{\mu c}\hat{\mathbf{s}}\mathbf{B}\Psi(t, \mathbf{x}), \end{aligned}$$

where $\Psi(t, \mathbf{x}) = (\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x})) \in \mathcal{E}^\otimes = \mathcal{E} \otimes \mathbb{C}^2$ and the additional factor 2 in the last term corresponds to (19.2).

Definition 19.3 i) For the Pauli Equation, the total angular momentum is the following mean value:

$$(19.9) \quad \mathbf{J} = \mathbf{L} + \mathbf{s}.$$

ii) The total magnetic moment is

$$(19.10) \quad \mathbf{m} = \frac{e}{2\mu c}\mathbf{L} + \frac{e}{\mu c}\mathbf{s}.$$

The unperturbed Pauli operator \mathcal{P} with $\mathbf{B} = 0$ commutes with all T_g and S_g . Hence, the orbital and spin momenta, \mathbf{L} and \mathbf{s} , are conserved quantities if $\mathbf{B} = 0$. Therefore, \mathbf{J} is also conserved if $\mathbf{B} = 0$.

Remark 19.4 For the Pauli equation with $\mathbf{B} = 0$ we have three conserved angular momenta \mathbf{L} , \mathbf{s} and \mathbf{J} . Hence, the identification of the total angular momentum with the vector \mathbf{J} , is not well justified.

For $\mathbf{B} = (0, 0, B) \neq 0$, the equation becomes

$$(19.11) \quad \begin{aligned} i\hbar\partial_t\Psi(t, \mathbf{x}) &= \mathcal{P}\Psi(t, \mathbf{x}) \\ &= -\frac{1}{2\mu}\hbar^2\Delta\Psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|)\Psi(t, \mathbf{x}) - \frac{e}{2\mu c}\hat{\mathbf{L}}_3B\Psi(t, \mathbf{x}) - \frac{e}{\mu c}\hat{\mathbf{s}}_3B\Psi(t, \mathbf{x}). \end{aligned}$$

Remark 19.5 *The operator \mathcal{P} commutes with $\hat{\mathbf{L}}^2 := \hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2$, $\hat{\mathbf{L}}_3$, $\hat{\mathbf{s}}^2 := \hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2$, $\hat{\mathbf{s}}_3$ and $\hat{\mathbf{J}}^2 := \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$, $\hat{\mathbf{J}}_3$. Therefore, the corresponding mean values \mathbf{L}^2 , \mathbf{L}_3 , $\mathbf{s}^2 = 3/4$, \mathbf{s}_3 and \mathbf{J}^2 , $\mathbf{J}_3 = \mathbf{L}_3 + \mathbf{s}_3$, are conserved.*

19.4 Pauli Equation. General Maxwell Field

The natural extension of the Pauli equation for a general Maxwell field reads

$$(19.12) \quad [i\hbar\partial_t - e\phi(t, \mathbf{x})]\Psi(t, \mathbf{x}) = \frac{1}{2\mu}[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\Psi(t, \mathbf{x}) - \frac{e}{\mu c}\hat{\mathbf{s}}\mathbf{B}(t, \mathbf{x})\Psi(t, \mathbf{x}).$$

It corresponds to the Lagrangian density (cf. (2.6))

$$(19.13) \quad \mathcal{L}_P(x, \Psi, \nabla\Psi) = [i\hbar\nabla_0 - e\phi(x)]\Psi \cdot \Psi - \frac{1}{2\mu}|[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)]\Psi|^2 - \frac{e}{\mu c}[\hat{\mathbf{s}}\mathbf{B}(x)\Psi] \cdot \Psi,$$

where $\mathbf{B}(x) = \text{rot } \mathbf{A}(x)$ and ' \cdot ' stands for the scalar product in $\mathbb{C}^2 \equiv \mathbb{R}^4$. This suggests the following extension to the Lagrangian density of the coupled Maxwell-Pauli equations (cf. (7.1))

$$(19.14) \quad \mathcal{L}_{MP} = [i\hbar\nabla_0 - e\phi(x)]\Psi \cdot \Psi - \frac{1}{2\mu}|[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)]\Psi|^2 - \frac{e}{\mu c}[\hat{\mathbf{s}}\mathbf{B}(x)\Psi] \cdot \Psi - \frac{1}{16\pi}\mathcal{F}^{\alpha\beta}\mathcal{F}_{\alpha\beta}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu\mathcal{A}^\nu - \partial^\nu\mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu$. The corresponding Euler-Lagrange equations are the coupled Maxwell-Pauli equations (cf. (7.2))

$$(19.15) \quad \left\{ \begin{aligned} [i\hbar\partial_t - e\phi(t, \mathbf{x})]\Psi(t, \mathbf{x}) &= \frac{1}{2\mu}[-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2\Psi(t, \mathbf{x}) - \frac{e}{\mu c}\hat{\mathbf{s}}\mathbf{B}(t, \mathbf{x})\Psi(t, \mathbf{x}). \\ \frac{1}{4\pi}\nabla_\alpha\mathcal{F}^{\alpha\beta}(t, \mathbf{x}) &= \begin{pmatrix} \rho := e|\Psi(t, \mathbf{x})|^2 \\ \frac{\mathbf{j}}{c} := \frac{e}{\mu c}[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]\Psi(t, \mathbf{x}) \cdot \Psi(t, \mathbf{x}) + \frac{e}{\mu c}\text{rot} \left([\hat{\mathbf{s}}\Psi(t, \mathbf{x})] \cdot \Psi(t, \mathbf{x}) \right) \end{pmatrix} \end{aligned} \right.$$

19.5 Application to the Stern-Gerlach Experiment

Theorem 16.2 allows us to construct quantum stationary states corresponding to the Pauli equation (19.11) describing the hydrogen atom. Let us define the vector-functions

$$(19.16) \quad \Psi_{lmn}^+ = \begin{pmatrix} \psi_{lmn} \\ 0 \end{pmatrix}, \quad \Psi_{lmn}^- = \begin{pmatrix} 0 \\ \psi_{lmn} \end{pmatrix}, \quad n = 1, 2, 3, \dots, \quad l \leq n-1, \quad m = -l, \dots, l,$$

where the functions ψ_{lmn} are given by (8.34). They are eigenfunctions of the operator \mathcal{P} corresponding to the energies

$$(19.17) \quad E_{mn}^\pm := -2\pi\hbar R/n^2 + m\hbar\omega_{\mathcal{L}} \pm \hbar\omega_{\mathcal{L}}.$$

In particular, the Schrödinger ground state ψ_{001} generates two stationary states Ψ_{001}^\pm with distinct energies $E_{01}^\pm := -2\pi\hbar R \pm \hbar\omega_{\mathcal{L}}$ and distinct spin magnetic moments $\pm \frac{e\hbar}{2\mu c}$. This explains the double splitting of the beam in the Stern-Gerlach experiment.

Remark 19.6 The double splitting of spectral lines is observed experimentally **even in the absence of the external magnetic field**, when formally $\omega_{\mathcal{L}} := eB/(2\mu c) = 0$. This means that empirically the Larmor frequency $\omega_{\mathcal{L}} \neq 0$ even for $B = 0$. It is natural to think that it is produced by the **intrinsic magnetic field** generated by the electric current in the coupled Maxwell-Pauli equations (19.15). Let us recall that the formulas (19.16), (19.17) are obtained in the *Born approximation* neglecting the intrinsic field. So the double splitting in the absence of the external magnetic field might be explained by the nonlinear self-interaction in the coupled Maxwell-Pauli equations. An alternative mechanism of generation of the magnetic field has been proposed by Thomas [97] and Frenkel [34] (see below).

Comment 19.7 The classical interpretation of the splitting term $\pm\hbar\omega_{\mathcal{L}}$ in (19.17) means that the "projection of the spin magnetic moment" to the axis Ox_3 is equal to $\pm\frac{e}{2\mu c}$. The magnitude of the projection is attained "instantly" with the magnetic field, which contradicts the classical picture (nonzero moment of inertia, etc). In the quantum context the instant reaction is not surprising in view of the conjecture **A** (see Preface). Namely, it should be interpreted as the change of the attractor of the system, and the formulas (19.17) determine the instant bifurcation of the (point) attractor.

20 Einstein-de Haas and Anomalous Zeemann Effects

We consider a modification of the Pauli equation taking into account the interaction between the orbital and spin angular momentum. The Russell-Saunders method gives a satisfactory explanation of the Einstein-de Haas and anomalous Zeemann effects.

The Pauli equation in the form (19.8) is not sufficient for the explanation of the anomalous Zeemann effect. This follows from the formulas (19.17), (19.17). Namely, the selection rules $m \mapsto m, m \pm 1$ give the splitting $\omega_0 \mapsto \omega_0, \omega_0 \pm \omega_L$ of the unperturbed spectral line ω_0 . Hence, the splitting gives the triplet corresponding to the normal Zeemann effect (cf. Remark 16.4).

This situation is related to the fact that the Pauli equation (19.8) does not take into account the interaction between the orbital and spin angular momenta (see Remark 19.4), while the interaction is suggested by the great success of the phenomenological *vector model* (see Exercise 15). That model suggests the interaction through the Maxwell field. Namely, the orbital angular momentum is related to the rotational motion of the electron, i.e., the circular current which generates the magnetic field acting on the spin magnetic moment. Similarly, the spin angular momentum is related to the ‘intrinsic rotation’ of the electron which generates the magnetic field acting on the orbital motion of the electron.

20.1 Spin-Orbital Coupling

To explain the anomalous Zeemann effect, we have to modify the equation (19.8) taking into account the interaction between the orbital and spin angular momenta. In spirit, the interaction might be described by the coupled Maxwell-Pauli equations (19.15) (see Remark 19.6). An alternative *relativistic* modification has been found by Thomas [97] and Frenkel [34]. Namely, the modification takes into account the magnetic field arising in the “moving frame of the electron”. The magnetic field is expressed in the electrostatic radial potential $\phi(|\mathbf{x}|)$ by the Lorentz formulas (23.17): in the first order approximation w.r.t. $\beta = v/c$, we get

$$(20.1) \quad \tilde{\mathbf{B}} = \frac{1}{c} \mathbf{E} \times \mathbf{v} = -\frac{1}{c} \nabla \phi(|\mathbf{x}|) \times \mathbf{v} = -\frac{1}{\mu c} \phi'(|\mathbf{x}|) \frac{\mathbf{x}}{|\mathbf{x}|} \times \mathbf{p} = -\frac{1}{\mu c |\mathbf{x}|} \phi'(|\mathbf{x}|) \hat{\mathbf{L}}$$

where we set *formally* $\mathbf{v} = \mathbf{p}/\mu$. This magnetic field produces the corresponding correction term $-\frac{e}{\mu c} \hat{\mathbf{s}} \tilde{\mathbf{B}} \Psi(t, \mathbf{x})$ in the RHS of (19.8). Furthermore, the Thomas and Frenkel phenomenological arguments imply an additional factor 1/2 which is justified by the relativistic Dirac theory [82]. Hence, finally, the modified equation reads

$$(20.2) \quad \begin{aligned} i\hbar \partial_t \Psi(t, \mathbf{x}) &= \mathcal{P}_m \Psi(t, \mathbf{x}) \\ &:= -\frac{1}{2\mu} \hbar^2 \Delta \Psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|) \Psi(t, \mathbf{x}) + \frac{e}{2\mu^2 c^2} \frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|} \hat{\mathbf{L}} \hat{\mathbf{s}} \Psi(t, \mathbf{x}) - \frac{e}{2\mu c} \hat{\mathbf{L}} \mathbf{B} \Psi(t, \mathbf{x}) - \frac{e}{\mu c} \hat{\mathbf{s}} \mathbf{B} \Psi(t, \mathbf{x}). \end{aligned}$$

In the case $B = 0$ the equation becomes

$$(20.3) \quad \begin{aligned} i\hbar \partial_t \Psi(t, \mathbf{x}) &= \overline{\mathcal{P}}_m \Psi(t, \mathbf{x}) \\ &:= -\frac{1}{2\mu} \hbar^2 \Delta \Psi(t, \mathbf{x}) + e\phi(|\mathbf{x}|) \Psi(t, \mathbf{x}) + \frac{e}{2\mu^2 c^2} \frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|} \hat{\mathbf{L}} \hat{\mathbf{s}} \Psi(t, \mathbf{x}). \end{aligned}$$

Remarks 20.1 *i)* The correction term $\sim \hat{\mathbf{L}} \hat{\mathbf{s}}$ is called the **Russell-Saunders spin-orbital coupling**.

ii) A "rigorous" justification for the correction term follows from the Dirac relativistic equation. Namely, just this term appears in an asymptotic expansion of the Dirac equation in the series of c^{-1} (see [89, Vol. II] and [9, 84]).

iii) The term is a genuine relativistic correction since it is provided by the Lorentz transformation (23.17) of the Maxwell field.

We will see that the stationary states and the energies for the equation (20.2) depend on the azimuthal quantum number (in contrast to (19.16), (19.17)) and describe perfectly the anomalous Zeemann effect as well as the Einstein-de Haas experiment. First let us assume that $\mathbf{B} = (0, 0, B)$ and rewrite (20.2) in the form

$$(20.4) \quad i\hbar\partial_t\Psi(t, \mathbf{x}) = \overline{\mathcal{P}}_m\Psi(t, \mathbf{x}) - \hat{\mathbf{m}}_3 B\Psi(t, \mathbf{x}),$$

where $\hat{\mathbf{m}} = \frac{e}{2\mu c}\hat{\mathbf{L}} + \frac{e}{\mu c}\hat{\mathbf{s}}$ in accordance with the definition (19.10). We will apply the perturbation theory to calculate the spectrum of the equation (20.4) for small $|B|$. More precisely, we will calculate the corrections to the eigenvalues of the unperturbed equation (20.4) with $B = 0$. This is sufficient since just these corrections describe the splitting of the spectral lines in the magnetic field.

We will also calculate the Landé factor for quantum stationary states of the equation (20.4). The factor explains the Einstein-de Haas experiment and anomalous Zeemann effect. By Definition (19.10), the magnetic moment for a state Ψ is defined as

$$(20.5) \quad \mathbf{m} = \langle \Psi, \hat{\mathbf{m}}\Psi \rangle.$$

The Landé factor g for a state Ψ is defined by

$$(20.6) \quad \frac{\mathbf{m}_3}{\mathbf{J}_3} = g\frac{e}{2\mu c},$$

and this factor (or rather its maximal value) is measured in the Einstein-de Haas experiment.

20.2 Russell-Saunders Quantum Numbers

The main point of the Russell-Saunders method is the analysis of the symmetry properties of the problem which correspond to the correction term $\hat{\mathbf{L}}\hat{\mathbf{s}}$. Namely, the operators $\mathcal{P}_m, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3 := \hat{\mathbf{L}}_3 + \hat{\mathbf{s}}_3$ commute with each other. On the other hand, the momenta $\hat{\mathbf{L}}_3$ and $\hat{\mathbf{s}}_3$ do not commute with $\hat{\mathbf{L}}\hat{\mathbf{s}}$ and hence with the generator \mathcal{P}_m . Hence, we cannot use the classification of type (19.16) for the stationary states by the 'quantum numbers' $n, l, m, \pm 1$ which correspond to the eigenvalues of the operators $\mathcal{P}, \hat{\mathbf{L}}^2, \hat{\mathbf{L}}_3, \hat{\mathbf{s}}_3$. Let us choose different quantum numbers.

Exercise 20.2 Check that the operators $\mathcal{P}_m, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3$ commute with each other. **Hints:**

i) (20.2) implies that

$$(20.7) \quad \mathcal{P}_m = -\frac{1}{2\mu}\hbar^2\Delta + e\phi(|\mathbf{x}|) + \frac{e}{2\mu^2c^2}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\hat{\mathbf{L}}\hat{\mathbf{s}} - \frac{e}{2\mu c}[\hat{\mathbf{L}}_3 + 2\hat{\mathbf{s}}_3]B.$$

ii) $\hat{\mathbf{L}}\hat{\mathbf{s}}$ commutes with $\hat{\mathbf{L}}^2$ since each $\hat{\mathbf{L}}_n$ and $\hat{\mathbf{s}}_n$ commutes with $\hat{\mathbf{L}}^2$.

iii) $\hat{\mathbf{J}}^2$ commutes with $\hat{\mathbf{L}}^2$ since $\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + 2\hat{\mathbf{L}}\hat{\mathbf{s}} + \hat{\mathbf{s}}^2$.

iv) $\hat{\mathbf{L}}\hat{\mathbf{s}}$ commutes with $\hat{\mathbf{J}}^2$ and $\hat{\mathbf{J}}_3$ since $2\hat{\mathbf{L}}\hat{\mathbf{s}} = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{s}}^2$.

Therefore, we could expect that the operators in the space $\mathcal{E}^\otimes = \mathcal{E} \otimes \mathbb{C}^2$ can be simultaneously diagonalized. Let us recall that possible eigenvalues of the operators are, respectively, $E, \hbar^2L(L+1), \hbar^2J(J+1)$ and $\hbar M$, where $J = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ and $M = -J, \dots, J$ (see Sections 10.2 and 10.3). Let us introduce Dirac's notation.

Definition 20.3 $|E, L, J, M\rangle$ denotes an eigenvector which corresponds to the eigenvalues $E, \hbar^2L(L+1), \hbar^2J(J+1), \hbar M$, of the operators $\mathcal{P}, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3$, respectively.

20.3 Landé Formula

Theorem 20.4 *The Landé factor $g = g(\Psi)$ corresponding to the stationary state $\Psi = |E, L, J, M\rangle$ is given by*

$$(20.8) \quad g = \frac{3}{2} + \frac{3/4 - L(L+1)}{2J(J+1)}.$$

Proof The vector Ψ belongs to one of the irreducible subspaces invariant with respect to the operators $\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2, \hat{\mathbf{J}}_3$. Indeed, $\hat{\mathbf{J}}^2\Psi = \hbar^2 J(J+1)\Psi$ and $\hat{\mathbf{J}}_3\Psi = \hbar M\Psi$. Therefore Ψ is an element e_M of the canonical basis e_{-J}, \dots, e_J by Proposition 9.8, and all basis vectors $e_{M'}$ are obtained from e_M by an application of the operators $\hat{\mathbf{J}}_{\pm} := \hat{\mathbf{J}}_1 \pm i\hat{\mathbf{J}}_2$.

For a linear operator A in \mathcal{E}^{\otimes} , let us denote the matrix element $A^{M, M'} = \langle Ae_M, e_{M'} \rangle$. Then (19.9) and (19.10) imply that

$$(20.9) \quad \mathbf{m}_3 = \hat{\mathbf{m}}_3^{M, M} = \frac{e}{2\mu c} (\hat{\mathbf{J}}_3^{M, M} + \hat{\mathbf{s}}_3^{M, M}) = \frac{e}{2\mu c} (\hbar M + \hat{\mathbf{s}}_3^{M, M}).$$

It remains to find the matrix element $\hat{\mathbf{s}}_3^{M, M}$ and calculate g from the definition (20.6). Let us collect the commutators of the operators:

$$(20.10) \quad \begin{cases} [\hat{\mathbf{J}}_1, \hat{\mathbf{s}}_1] = 0, & [\hat{\mathbf{J}}_1, \hat{\mathbf{s}}_2] = i\hbar\hat{\mathbf{s}}_3, & [\hat{\mathbf{J}}_1, \hat{\mathbf{s}}_3] = -i\hbar\hat{\mathbf{s}}_2, \\ [\hat{\mathbf{J}}_2, \hat{\mathbf{s}}_2] = 0, & [\hat{\mathbf{J}}_2, \hat{\mathbf{s}}_3] = i\hbar\hat{\mathbf{s}}_1, & [\hat{\mathbf{J}}_2, \hat{\mathbf{s}}_1] = -i\hbar\hat{\mathbf{s}}_3, \\ [\hat{\mathbf{J}}_3, \hat{\mathbf{s}}_3] = 0, & [\hat{\mathbf{J}}_3, \hat{\mathbf{s}}_1] = i\hbar\hat{\mathbf{s}}_2, & [\hat{\mathbf{J}}_3, \hat{\mathbf{s}}_2] = -i\hbar\hat{\mathbf{s}}_1, \end{cases}$$

where the second and the third line follow from the first one by cyclic permutations. Let us use the commutators (20.10) to calculate the commutators of the operators $\hat{\mathbf{J}}_{\pm}$ with $\hat{\mathbf{s}}_{\pm} := \hat{\mathbf{s}}_1 \pm i\hat{\mathbf{s}}_2$. We find

$$(20.11) \quad [\hat{\mathbf{J}}_-, \hat{\mathbf{s}}_+] = -2\hbar\hat{\mathbf{s}}_3, \quad [\hat{\mathbf{J}}_+, \hat{\mathbf{s}}_+] = 0.$$

The first formula implies the identity

$$(20.12) \quad \hat{\mathbf{J}}_-^{M, M+1} \hat{\mathbf{s}}_+^{M+1, M} - \hat{\mathbf{s}}_+^{M, M-1} \hat{\mathbf{J}}_-^{M-1, M} = -2\hbar\hat{\mathbf{s}}_3^{M, M},$$

since all other matrix elements $\hat{\mathbf{J}}_-^{M', M''} = 0$ for $M'' \neq M' + 1$ by (9.10). The non-zero matrix elements of $\hat{\mathbf{J}}_{\pm}$ are known from (9.10) and (9.7):

$$(20.13) \quad \hat{\mathbf{J}}_+^{M+1, M} = \hbar\sqrt{(J-M)(J+M+1)}, \quad \hat{\mathbf{J}}_-^{M-1, M} = \hbar\sqrt{(J+M)(J-M+1)}.$$

On the other hand, $\hat{\mathbf{s}}_+$ can be derived from the second identity (20.11). Taking the matrix element $(\cdot)^{M+1, M-1}$, we derive

$$(20.14) \quad \hat{\mathbf{J}}_+^{M+1, M} \hat{\mathbf{s}}_+^{M, M-1} - \hat{\mathbf{s}}_+^{M+1, M} \hat{\mathbf{J}}_+^{M, M-1} = 0.$$

Then (20.13) implies that

$$(20.15) \quad \frac{\hat{\mathbf{s}}_+^{M+1, M}}{\sqrt{(J-M)(J+M+1)}} = \frac{\hat{\mathbf{s}}_+^{M, M-1}}{\sqrt{(J-M+1)(J+M)}} =: A.$$

Substituting this into (20.12) and using the matrix elements of $\hat{\mathbf{J}}_-$ from (20.13), we get

$$(20.16) \quad \hat{\mathbf{s}}_3^{M, M} = AM.$$

It remains to calculate the constant A . We start with the identity

$$(20.17) \quad \hat{\mathbf{J}}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{s}})^2 = \hat{\mathbf{L}}^2 + 2\hat{\mathbf{L}}\hat{\mathbf{s}} + \hat{\mathbf{s}}^2 = \hat{\mathbf{L}}^2 + 2\hat{\mathbf{J}}\hat{\mathbf{s}} - \hat{\mathbf{s}}^2.$$

Hence

$$(20.18) \quad (\hat{\mathbf{J}}\hat{\mathbf{s}})^{M,M} = \hbar^2 \frac{J(J+1) - L(L+1) + 3/4}{2}$$

since $\hat{\mathbf{s}}^2 = 3/4$. On the other hand, the same matrix element can be expressed with the help of another identity

$$(20.19) \quad 2\hat{\mathbf{J}}\hat{\mathbf{s}} = \hat{\mathbf{s}}_+\hat{\mathbf{J}}_- + \hat{\mathbf{s}}_-\hat{\mathbf{J}}_+ + \hat{\mathbf{J}}_3\hat{\mathbf{s}}_3$$

as

$$(20.20) \quad (\hat{\mathbf{J}}\hat{\mathbf{s}})^{M,M} = \frac{1}{2}\hat{\mathbf{s}}_+^{M,M-1}\hat{\mathbf{J}}_-^{M-1,M} + \frac{1}{2}\hat{\mathbf{s}}_-^{M,M+1}\hat{\mathbf{J}}_+^{M+1,M} + \hbar M\hat{\mathbf{s}}_3^{M,M}.$$

Note that the matrix elements $\hat{\mathbf{s}}_+^{M+1,M}$ and $\hat{\mathbf{s}}_-^{M,M+1}$ are complex conjugate. Therefore, we have by (20.15) that

$$(20.21) \quad \hat{\mathbf{s}}_+^{M+1,M} = A\sqrt{(J-M)(J+M+1)} = \hat{\mathbf{s}}_-^{M,M+1},$$

since the constant A is real by (20.16). Finally, let us substitute (20.21), (20.18), (20.16) and (20.13) into (20.15). Then we get an equation for A which implies that

$$(20.22) \quad A = \hbar \frac{J(J+1) - L(L+1) + 3/4}{2J(J+1)}.$$

At last, (20.16) and (20.9) give that

$$(20.23) \quad \mathbf{m}_3 = \frac{e}{2\mu c}(\hbar M + AM) = \frac{e}{2\mu c}(1 + A/\hbar)\hbar M = \frac{e}{2\mu c}(1 + A/\hbar)\mathbf{J}_3.$$

Substituting (20.22) for A , we get (20.6), where g coincides with (20.8). ■

Remarks 20.5 *i) Our proof follows the calculations in [71].*

ii) The formula (20.8) has been obtained by Landé, [70], using the phenomenological vector model and the Bohr correspondence principle (see also [12] and [89, Vol.I]).

20.4 Application to the Einstein-de Haas and Anomalous Zeemann Effects

The Landé formula (20.8) explains satisfactorily the Einstein-de Haas experiments (see [12] and [89, Vol.I]). Furthermore, by *perturbation theory* (see [102]), the Landé factor (20.8) determines the splitting of the eigenvalues for (20.4) in the magnetic field $(0, 0, B)$ with small $|B|$. The resulting correction to the frequency $\omega = E/\hbar$ of the eigenstate $|E, L, J, M\rangle$ is equal to

$$(20.24) \quad \Delta\omega_{L,J,M} = -\hat{\mathbf{m}}_3^{M,M}B/\hbar = -\mathbf{m}_3B/\hbar = -g\frac{e}{2\mu c}MB = -g\omega_{\mathcal{L}}M$$

up to corrections $\mathcal{O}(B^2)$. This formula, together with the selection rules $J \mapsto J' = J \pm 1$, $M \mapsto M' = M, M \pm 1$ (suggested by the Bohr correspondence principle [12] and [89, Vol.I]), explains the multiplet structure in the anomalous Zeemann effect. Namely, the magnitude of the splitting (20.24), $g\frac{e}{2\mu c}\hbar B$, now depends on the quantum numbers L, J in contrast to (19.17) and (16.15). This explains the multiplet structure of the spectrum since (16.17) now becomes

$$(20.25) \quad \omega_{kk'} = \omega_{kk'}^0 - \omega_{\mathcal{L}}[g(k)M - g(k')M'] = \omega_{kk'}^0 - \omega_{\mathcal{L}}g(k)[M - M'] - \omega_{\mathcal{L}}[g(k) - g(k')]M',$$

where $k = (E, L, J, M)$, $k' = (E', L', J', M')$ and $\omega_{kk'}^0$ is the 'unperturbed' spectral line corresponding to $B = 0$. Hence, the radiation in the allowed transition does not reduce to the normal triplet (16.17).

The calculations are in excellent agreement with experimental observations (see [9, 12, 17] and [89, Vol.I]). This agreement was one of the greatest successes of quantum theory.

Remark 20.6 *The calculation of the splitting (20.25) does not depend on the magnitude of the unperturbed spectral line $\omega_{kk'}^0$. The magnitude can be calculated also with the help of the perturbation theory (see [9, 17] and [89, Vol.II]).*

Part V

Special Relativity

21 Electromagnetic Nature of Light

We deduce the Maxwell equations from the Coulomb and Biot-Savart-Laplace laws for the interaction of charges and currents. We also discuss connections of the Maxwell equations to the Einstein ideas on relativity, in particular, to the Lorentz transformation.

21.1 Maxwell Equations. Empirical Synthesis

Maxwell stated the equations written in the MKS rationalized (or SI) system of units:

$$(21.1) \quad \begin{cases} \operatorname{div} \mathbf{E}(t, \mathbf{x}) = \frac{1}{\varepsilon_0} \rho(t, \mathbf{x}), & \operatorname{rot} \mathbf{E}(t, \mathbf{x}) = -\dot{\mathbf{B}}(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0, & \frac{1}{\mu_0} \operatorname{rot} \mathbf{B}(t, \mathbf{x}) = \mathbf{j}(t, \mathbf{x}) + \varepsilon_0 \dot{\mathbf{E}}(t, \mathbf{x}). \end{cases}$$

The equations contain the dielectric permittivity and magnetic permeability of the vacuum, ε_0 and μ_0 , and do not contain explicitly the speed of light c . Maxwell deduced the equations from the Coulomb and Biot-Savart-Laplace laws for the interaction of charges and currents. The Coulomb law states the force of an electrostatic interaction of two charges $q_{1,2}$ at a distance r :

$$(21.2) \quad \mathbf{F}_2 = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2 \mathbf{e}_{1,2}}{r^2},$$

where $\mathbf{e}_{1,2}$ is the unit vector directed from the charge at q_1 to the charge at q_2 , and \mathbf{F}_2 is the force acting onto the second elementary charge. Similarly, the Biot-Savart-Laplace law states the force of a magnetic interaction of two (stationary) elementary currents $I_k d\mathbf{l}_k$, $k = 1, 2$, at a distance r :

$$(21.3) \quad \mathbf{F}_2 = \frac{\mu_0}{4\pi} \frac{I_2 d\mathbf{l}_2 \times (I_1 d\mathbf{l}_1 \times \mathbf{e}_{1,2})}{r^2},$$

where \mathbf{F}_2 is the force acting onto the second elementary current $I_2 d\mathbf{l}_2$. The values of ε_0 and μ_0 were measured in a laboratory for electromagnetic experiments with a high accuracy. In the MKS system

$$(21.4) \quad \varepsilon_0 \approx \frac{1}{4\pi \cdot 9 \cdot 10^9} \frac{As}{Vm}, \quad \mu_0 \approx 4\pi 10^{-7} \frac{Vs}{Am}.$$

The first equation

First, (21.2) implies that the electric field of the first elementary charge is

$$(21.5) \quad \mathbf{E}(\mathbf{x}_2) := \frac{\mathbf{F}_2}{q_2} = \frac{1}{4\pi\varepsilon_0} \frac{q_1 \mathbf{e}_{1,2}}{|\mathbf{x}_{1,2}|^2},$$

where $\mathbf{x}_{1,2} := \mathbf{x}_2 - \mathbf{x}_1$ and $\mathbf{x}_1, \mathbf{x}_2$ are the vector-positions of the charges. Then for a distribution of the charges, $\rho(\mathbf{x}_1) d\mathbf{x}_1$, we obtain by the principle of superposition,

$$(21.6) \quad \mathbf{E}(\mathbf{x}_2) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\mathbf{x}_1) \mathbf{e}_{1,2}}{|\mathbf{x}_{1,2}|^2} d\mathbf{x}_1,$$

Differentiation gives

$$(21.7) \quad \operatorname{div} \mathbf{E}(\mathbf{x}_2) = \frac{1}{\varepsilon_0} \rho(\mathbf{x}_2),$$

which coincides with the first equation of (21.1).

Exercise 21.1 Check (21.7) for $\rho(\mathbf{x}) \in C_0^\infty(\mathbb{R}^3)$. **Hint:** use that $\operatorname{div}_{\mathbf{x}_2} \frac{\mathbf{e}_{1,2}}{|\mathbf{x}_{1,2}|^2} = \Delta_{\mathbf{x}_2} \frac{1}{|\mathbf{x}_{1,2}|} = -4\pi\delta(\mathbf{x}_{1,2})$.

Last equation for stationary currents

Second, (21.3) means, by Ampere's law, that the magnetic field of the first elementary current is

$$(21.8) \quad \mathbf{B}(\mathbf{x}_2) = \frac{\mu_0}{4\pi} \frac{I_1 d\mathbf{l}_1 \times \mathbf{e}_{1,2}}{|\mathbf{x}_{1,2}|^2}.$$

Integrating, we obtain by the principle of superposition,

$$(21.9) \quad \mathbf{B}(\mathbf{x}_2) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{x}_1) \times \mathbf{e}_{1,2}}{|\mathbf{x}_{1,2}|^2} d\mathbf{x}_1,$$

where $\mathbf{j}(\mathbf{x}_1)$ is the current density at the point \mathbf{x}_1 . Differentiating, we get

$$(21.10) \quad \text{rot } \mathbf{B}(\mathbf{x}_2) = \mu_0 \mathbf{j}(\mathbf{x}_2),$$

which coincides with the last equation of (21.1 in the case of the **stationary currents**.

Exercise 21.2 Check (21.9) for $\mathbf{j}(\mathbf{x}) \in C_0^\infty(\mathbb{R}^3)$. **Hint:** use the charge continuity equation $\text{div } \mathbf{j}(\mathbf{x}) \equiv 0$ for the stationary currents.

Last equation for nonstationary currents: Maxwell's Displacement Current

The divergence of the LHS of the last Maxwell equation vanishes since $\text{div rot } = 0$. On the other hand, the divergence of the RHS, $\text{div } \mathbf{j}(t, \mathbf{x})$ generally does not vanish for nonstationary currents. To solve this contradiction, Maxwell has completed the last equation by the *displacement current* $\varepsilon_0 \dot{\mathbf{E}}(t, \mathbf{x})$. The current provides a vanishing divergence of the RHS:

$$(21.11) \quad \text{div } \mathbf{j}(t, \mathbf{x}) + \text{div } \varepsilon_0 \dot{\mathbf{E}}(t, \mathbf{x}) = 0$$

by the charge continuity equation $\text{div } \mathbf{j}(t, \mathbf{x}) + \dot{\rho}(t, \mathbf{x}) = 0$ and the first Maxwell equation.

Faraday's equations

The second and third Maxwell equations have been divined and checked experimentally by Faraday.

21.2 Equations for Potentials

Let us repeat the introduction of the Maxwell potentials (see Lecture 4) in the MKS system. Namely, $\text{div } \mathbf{B}(t, \mathbf{x}) = 0$ implies that $\mathbf{B}(t, \mathbf{x}) = \text{rot } \mathbf{A}(t, \mathbf{x})$. Then $\text{rot } \mathbf{E}(t, \mathbf{x}) = -\dot{\mathbf{B}}(t, \mathbf{x})$ implies $\text{rot } [\mathbf{E}(t, \mathbf{x}) + \dot{\mathbf{A}}(t, \mathbf{x})] = 0$ hence $\mathbf{E}(t, \mathbf{x}) + \dot{\mathbf{A}}(t, \mathbf{x}) = -\nabla_{\mathbf{x}}\phi(t, \mathbf{x})$. Finally,

$$(21.12) \quad \mathbf{B}(t, \mathbf{x}) = \text{rot } \mathbf{A}(t, \mathbf{x}), \quad \mathbf{E}(t, \mathbf{x}) = -\nabla_{\mathbf{x}}\phi(t, \mathbf{x}) - \dot{\mathbf{A}}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Exercise 21.3 Prove the existence of the potentials (21.12). **Hint:** Use the Fourier transform.

The choice of the potentials is not unique since the *gauge transformation*

$$(21.13) \quad \phi(t, \mathbf{x}) \mapsto \phi(t, \mathbf{x}) + \dot{\chi}(t, \mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) \mapsto \mathbf{A}(t, \mathbf{x}) - \nabla_{\mathbf{x}}\chi(t, \mathbf{x})$$

does not change the fields $\mathbf{E}(t, \mathbf{x})$, $\mathbf{B}(t, \mathbf{x})$ for any function $\chi(t, \mathbf{x}) \in C^1(\mathbb{R}^4)$. Therefore, it is possible to satisfy an additional *gauge condition*. Let us choose for example the *Lorentz gauge*

$$(21.14) \quad \varepsilon_0 \dot{\phi}(t, \mathbf{x}) + \frac{1}{\mu_0} \text{div } \mathbf{A}(t, \mathbf{x}) = 0, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Let us express the Maxwell equations (21.1) in terms of the potentials. Substitution of (21.12) into the first Maxwell equation leads to $\frac{4\pi}{\varepsilon_0}\rho(t, \mathbf{x}) = \operatorname{div} \mathbf{E}(t, \mathbf{x}) = -\Delta\phi(t, \mathbf{x}) - \operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x})$. Eliminating $\operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x})$ by the differentiation of (21.14) in time, $\varepsilon_0\ddot{\phi}(t, \mathbf{x}) + \frac{1}{\mu_0}\operatorname{div} \dot{\mathbf{A}}(t, \mathbf{x}) = 0$, we get

$$(21.15) \quad [\varepsilon_0\mu_0\partial_t^2 - \Delta]\phi(t, \mathbf{x}) = \frac{1}{\varepsilon_0}\rho(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Similarly, substituting (21.12) into the last Maxwell equation, we get

$$(21.16) \quad \frac{1}{\mu_0}\operatorname{rot} \operatorname{rot} \mathbf{A}(t, \mathbf{x}) = \mathbf{j}(t, \mathbf{x}) + \varepsilon_0\dot{\mathbf{E}}(t, \mathbf{x}) = \mathbf{j}(t, \mathbf{x}) - \varepsilon_0\nabla_{\mathbf{x}}\dot{\phi}(t, \mathbf{x}) - \varepsilon_0\ddot{\mathbf{A}}(t, \mathbf{x}).$$

Exercise 21.4 *Prove the identity*

$$(21.17) \quad \operatorname{rot} \operatorname{rot} = -\Delta + \nabla_{\mathbf{x}}\operatorname{div} .$$

Substituting (21.17) to (21.16) and eliminating $\nabla_{\mathbf{x}}\dot{\phi}(t, \mathbf{x})$ by application of $\nabla_{\mathbf{x}}$ to (21.14), we get

$$(21.18) \quad [\varepsilon_0\mu_0\partial_t^2 - \Delta]\mathbf{A}(t, \mathbf{x}) = \mu_0\mathbf{j}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

Remark 21.5 *The arguments above show that the Maxwell equations (21.1) are equivalent to the system of two wave equations (21.15), (21.18) for the potentials with the Lorentz gauge condition (21.14).*

Maxwell deduced the wave equations (21.15) and (21.18) with the coefficient $\varepsilon_0\mu_0$ and at that time it was known that this coefficient is equal to $1/v^2$, where v is the propagation velocity of the solutions. He calculated $v = 1/\sqrt{\varepsilon_0\mu_0}$ using the values (21.4). The values (21.4) give $v \approx 3 \cdot 10^8 \text{ m/s}$ which almost coincides with the speed of light c . Hence the equations (21.15) and (21.18) become

$$(21.19) \quad \square\phi(t, \mathbf{x}) = \frac{1}{\varepsilon_0}\rho(t, \mathbf{x}), \quad \square\mathbf{A}(t, \mathbf{x}) = \mu_0\mathbf{j}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

where $\square := \frac{1}{c^2}\partial_t^2 - \Delta$. Furthermore, Maxwell found that the electromagnetic waves are transversal like the light waves. This is why Maxwell suggested to identify the electromagnetic waves with light. Below we will use *Gaussian units* and write (21.19) as

$$(21.20) \quad \square\phi(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}), \quad \square\mathbf{A}(t, \mathbf{x}) = \frac{4\pi}{c}\mathbf{j}(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

21.3 Problem of Luminiferous Ether: Michelson and Morley Experiment

The discovery of Maxwell led to a new very difficult question. The propagation velocity can be equal to c only in a **unique preferred reference frame** in which the Maxwell equations have the form (21.1) or (21.20). The preferred reference frame is called the frame of the *luminiferous ether*. In all other frames, the propagation velocity depends on a direction and vector of the relative velocity of the frame of reference, if space-time is transformed by the classical *Galilean transformations*

$$(21.21) \quad \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} t' \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} t \\ x_1 - vt \\ x_2 \\ x_3 \end{pmatrix}.$$

Therefore, the Maxwell equations are not invariant with respect to the Galilean transformations. This is why Michelson and Morley started around 1880 the famous experiment to identify the preferred *luminiferous ether frame* with the frame in which the Sun is at rest. They have tried to check that the Earth moves with respect to the luminiferous ether, i.e. the velocity of light along and against the velocity of the Earth differs by twice the velocity of the Earth. Concretely, they compared the wavelengths of light along and against the velocity of the Earth motion around the Sun. However, the result (1887) was negative and very discouraging: the wave lengths were identical with a high accuracy, hence the propagation velocity does not depend on the frame of reference! Astronomical observations of *double stars* by de Sitter (1908) confirmed the negative result of Michelson and Morley. Also the experiment of Trouton and Noble confirmed the negative result.

21.4 Time in a Moving Frame: Lorentz transformations

Various partial explanations of the negative results were proposed by Ritz, Fitzgerald, Lorentz and others. The complete explanation was provided in 1905 by Einstein, who was able to cumulate the Maxwell and Lorentz ideas into a new complete theory. The main novelty was the following postulate of the Einstein theory:

The time in a moving frame is distinct from the time in the rest frame!

We will prove that the transformation of space-time coordinates from the rest frame to the moving frame of reference is given by the **Lorentz formulas**

$$(21.22) \quad \begin{pmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} ct' \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \frac{ct - \beta x_1}{\sqrt{1 - \beta^2}} \\ \frac{x_1 - \beta ct}{\sqrt{1 - \beta^2}} \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1 - \beta^2}} & -\frac{\beta}{\sqrt{1 - \beta^2}} & 0 & 0 \\ -\frac{\beta}{\sqrt{1 - \beta^2}} & \frac{1}{\sqrt{1 - \beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

where (t, x_1, x_2, x_3) stands for the time-space coordinates in the rest frame and (t', x'_1, x'_2, x'_3) corresponds to the moving frame if the relative velocity is $(v, 0, 0)$, $|v| < c$ and $\beta := v/c$.

Exercise 21.6 Check that the wave equation

$$(21.23) \quad \left[\frac{1}{c^2} \partial_t^2 - \Delta \right] \phi(t, \mathbf{x}) = f(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \mathbb{R}^4$$

is invariant with respect to the transformations (21.22).

Hints: i) Set $c = 1$ and use that $\partial_2 = \partial'_2$, $\partial_3 = \partial'_3$. ii) Check that the 1D equation $(\partial_t^2 - \partial_1^2)\phi(t, x) = f(t, x)$, $(t, x) \in \mathbb{R}^2$ is equivalent to $[(\partial'_t)^2 - (\partial'_1)^2]\phi'(t', x') = f'(t', x')$ if $t' = at - bx$ and $x' = ax - bt$ with $a^2 - b^2 = 1$. By definition, $\phi'(t', x') = \phi(t, x)$ and $f'(t', x') = f(t, x)$.

Exercise 21.7 Check that the wave equation (21.23) is **not** invariant with respect to the **Galilean transformation** (21.21). **Hint:** In the new variables, the equation (21.23) becomes

$$(21.24) \quad \left[\frac{1}{c^2} (\partial_{t'} - v \partial_{x'})^2 - \Delta_{x'} \right] \phi'(t', \mathbf{x}') = f'(t', \mathbf{x}'), \quad (t', \mathbf{x}') \in \mathbb{R}^4.$$

Remark 21.8 For small velocities, $|v| \ll c$, the Galilean transformation (21.21) is close to the Lorentz one, (21.22), and the coefficients of the equation (21.24) are close to the ones of (21.23).

22 The Einstein Special Relativity and Lorentz Group

The formulas (21.22) define a particular transformation of the *Lorentz group*. We define the Lorentz group as transformations of the space-time preserving the form of the wave equation. This property follows from the Einstein's postulate of Special Relativity Theory extending the *Galilei's principle of relativity* from mechanics to electrodynamics. Namely, the main point of Einstein's *special relativity* is the following postulate:

E: Maxwell Equations have an identical form in all Inertial Frames

The postulate implies also an invariance of the wave equation (21.23) in the case $f(x) = 0$. Let us calculate all possible transformations which leave the equation with $f(x) \neq 0$ invariant.

First, introduce a new time variable $x_0 := ct$ and rewrite the wave equation (21.23) in the form

$$(22.1) \quad \square\phi(x) := g^{\alpha\beta}\partial_\alpha\partial_\beta\phi(x) = f(x), \quad x \in \mathbb{R}^4,$$

where $g^{\alpha,\beta} = \text{diag}(1, -1, -1, -1)$, $\partial_\alpha := \frac{\partial}{\partial x_\alpha}$ and the summation in the repeated indexes is understood, $\alpha, \beta = 0, 1, 2, 3$.

We look for maps $\Lambda : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ such that the equation keeps its form in new variables $x' = \Lambda x$, i.e., (22.1) is equivalent to

$$(22.2) \quad \square\phi'(x') = g_{\alpha'\beta'}\partial_{\alpha'}\partial_{\beta'}\phi'(x') = f'(x'), \quad x' \in \mathbb{R}^4,$$

where, by definition, $\phi'(x') := \phi(x)$ and $f'(x') := f(x)$. The map Λ , naturally, i) is continuous and moreover, ii) $\Lambda(x+x) = \Lambda x + \Lambda x$, $x \in \mathbb{R}^4$, which expresses the fact that space-time is homogeneous. Let us assume, additionally, the normalization condition iii) $\Lambda(0) = 0$. The three conditions imply necessarily that Λ is a *linear transformation* of the space \mathbb{R}^4 defined by a matrix:

$$(22.3) \quad x'_{\alpha'} = (\Lambda x)_{\alpha'} = \Lambda_{\alpha'}^\alpha x_\alpha, \quad \alpha' = 0, \dots, 3.$$

Exercise 22.1 Check that the conditions i), ii), iii), imply (22.3). **Hint:** First, check the corresponding one-dimensional statement with $x \in \mathbb{R}$.

Finally, rewrite the equation (22.1) in new variables (22.3). By the chain rule

$$(22.4) \quad \partial_\alpha = \frac{\partial x'_{\alpha'}}{\partial x_\alpha} \partial_{\alpha'} = \Lambda_{\alpha'}^\alpha \partial_{\alpha'}.$$

Hence, the equation in the new variables becomes

$$(22.5) \quad g_{\alpha,\beta} \Lambda_{\alpha'}^\alpha \Lambda_{\beta'}^\beta \partial_{\alpha'} \partial_{\beta'} \phi'(x') = f'(x'), \quad x' \in \mathbb{R}^4.$$

Comparing with (22.2), we get the system of algebraic equations

$$(22.6) \quad g_{\alpha,\beta} \Lambda_{\alpha'}^\alpha \Lambda_{\beta'}^\beta = g_{\alpha'\beta'}, \quad \alpha', \beta' = 0, 1, 2, 3.$$

In matrix form, the system is equivalent to the equation

$$(22.7) \quad \Lambda g \Lambda^t = g,$$

where Λ^t stands for the transposed matrix of Λ .

Exercise 22.2 Check that $|\det \Lambda| = 1$. **Hint:** take the determinant of both sides of (22.7).

Hence Λ is invertible and (22.7) is also equivalent to

$$(22.8) \quad (\Lambda^t)^{-1}g\Lambda^{-1} = g,$$

The matrix equation is equivalent to the invariance, with respect to the map Λ^{-1} , of the quadratic form

$$(22.9) \quad g(x, x) := (x, gx) = g_{\alpha\beta}x_\alpha x_\beta, \quad x \in \mathbb{R}^4$$

which is called the **Lorentz interval**. Then this form is invariant also with respect to the map Λ , hence

$$(22.10) \quad \Lambda^t g \Lambda = g,$$

or equivalently,

$$(22.11) \quad g_{\alpha\beta} \Lambda_\alpha^{\alpha'} \Lambda_\beta^{\beta'} = g_{\alpha'\beta'}, \quad \alpha', \beta' = 0, 1, 2, 3.$$

Definition 22.3 L is the set of all linear maps $\Lambda : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ satisfying (22.10).

Exercise 22.4 Check that the set L is a group.

Definition 22.5 i) Minkowski space is the space \mathbb{R}^4 endowed with the quadratic form (22.9).

ii) L is called the **Lorentz group**.

Example 22.6 The simplest example of a Lorentz transform is given by the matrices

$$(22.12) \quad \Lambda = \hat{R} := \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix},$$

where $R \in SO(3)$ is a rotation of the 3D space.

Remark 22.7 The wave equation (21.23) is invariant with respect to the transformation $x' = \hat{R}x$ since the Laplacian Δ is invariant with respect to the rotations.

Exercise 22.8 Check (22.7) for the matrix (22.12).

Exercise 22.9 Check that the map $R \mapsto \hat{R}$ is an isomorphism of the rotation group $SO(3)$ onto a subgroup of the Lorentz group L .

Exercise 22.10 Construct all Lorentz transformations of the form

$$(22.13) \quad \Lambda = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

Solution: (22.7) is equivalent to the matrix equation

$$(22.14) \quad \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which is equivalent to a system $a^2 - c^2 = 1$, $b^2 - d^2 = -1$, $ab - cd = 0$. Then $a = \pm \cosh \varphi$, $c = \pm \sinh \varphi$, $d = \pm \cosh \chi$, $b = \sinh \chi$, and $\pm \cosh \varphi \sinh \chi \pm \sinh \varphi \cosh \chi = 0$. Therefore, $\tanh \varphi = \pm \tanh \chi$ or $\varphi = \pm \chi$. Finally, we get four one-parametric families of the matrices

$$(22.15) \quad \Lambda_\pm^+ := \begin{pmatrix} \cosh \varphi & \sinh \varphi & 0 & 0 \\ \pm \sinh \varphi & \pm \cosh \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Lambda_\pm^- := - \begin{pmatrix} \cosh \varphi & \sinh \varphi & 0 & 0 \\ \pm \sinh \varphi & \pm \cosh \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

which are **hyperbolic rotations** (or **boosts**) in the angle φ in the plane x_0, x_1 .

Remark 22.11 *The Lorentz formulas (21.22) correspond to the matrix $\Lambda_{\dagger}^{\dagger}$ from (22.15) with $\cosh \varphi = 1/\sqrt{1 - \beta^2}$ and $\sinh \varphi = -\beta/\sqrt{1 - \beta^2}$, so $\tanh \varphi = -\beta$.*

23 Covariant Electrodynamics

Here we complete a justification of the Einstein postulate **E**. Namely, we still have to determine the transformation of the Maxwell Field by the Lorentz group.

Let us introduce the 4D fields and currents

$$(23.1) \quad \mathcal{A}(x) = (\phi(t, \mathbf{x}), \mathbf{A}(x)), \quad \mathcal{J}(x) = (\rho(x), \frac{1}{c} \mathbf{j}(x)), \quad x \in \mathbb{R}^4.$$

In this notation, the Maxwell equations (21.20) become

$$(23.2) \quad \square \mathcal{A}(x) = 4\pi \mathcal{J}(x), \quad x \in \mathbb{R}^4.$$

Hence, the transformations for the potentials $\mathcal{A}^\mu(x)$ and the currents $\mathcal{J}^\mu(x)$ must be identical since the wave operator \square is invariant with respect to the Lorentz group. We will prove below that the convective currents are transformed, like the 4-vector x , by the matrix Λ . Hence, the same holds for the potentials and all currents, so

$$(23.3) \quad \mathcal{A}'(x') = \Lambda \mathcal{A}(x), \quad \mathcal{J}'(x') = \Lambda \mathcal{J}(x), \quad x' = \Lambda x.$$

The Lagrangian approach gives another confirmation for (23.3). Namely, in Lecture 5 we have shown that the Maxwell equations are the canonical Euler-Lagrange equations corresponding to the Lagrangian density

$$(23.4) \quad \mathcal{L}(x, \mathcal{A}, \nabla \mathcal{A}) = -\frac{1}{16\pi} \mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu} - g(\mathcal{J}, \mathcal{A}), \quad (x, \mathcal{A}, \nabla \mathcal{A}) \in \mathbb{R}^4 \times \mathbb{R}^4 \times \mathbb{R}^{16}.$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$, $\mathcal{F}_{\mu\nu} := g_{\mu\mu'} g_{\nu\nu'} \mathcal{F}^{\mu'\nu'}$ and

$$(23.5) \quad g(\mathcal{J}, \mathcal{A}) := g_{\mu\nu} \mathcal{J}^\mu \mathcal{A}^\nu.$$

The Einstein postulate will hold if the density is invariant under the corresponding transformations for the fields. It is easy to check that the density is invariant if the potentials $\mathcal{A}^\mu(x)$ and the currents $\mathcal{J}^\mu(x)$ are transformed by (23.3).

Exercise 23.1 Check that the density (23.4) is invariant under the transformations (23.3). **Hint:** The bilinear form (23.5) is symmetric and invariant since the corresponding quadratic form is invariant.

At last, let us check that the transformation law (23.3) holds for the convective currents. It suffices to consider only point charges since a general convective current is a combination of moving point charges. Let $\mathbf{x} = \mathbf{x}(x_0)$ be a C^1 trajectory of a point charge, so its 4-current is given by

$$(23.6) \quad \mathcal{J}^0(x) = e\delta(\mathbf{x} - \mathbf{x}(x_0)), \quad \mathcal{J}^k(x) = e\mathbf{v}_k(x_0)\delta(\mathbf{x} - \mathbf{x}(x_0)),$$

where $\mathbf{v}_k(x_0) := \frac{d\mathbf{x}_k(x_0)}{dx_0}$.

Remark 23.2 The expressions (23.6) for the convective current are equivalent to the postulating of the charge-invariance under the motion which is referred traditionally as to an experimental fact (“charge invariance”, [46, 53]).

Proposition 23.3 The function (23.6) is Lorentz-covariant, i.e. for every Lorentz transformation $\Lambda \in \mathbf{L}$,

$$(23.7) \quad \mathcal{J}'(x') = \Lambda \mathcal{J}(x), \quad x' = \Lambda x.$$

Proof Since $\mathcal{J}(x)$ is a distribution, (23.7) means that

$$(23.8) \quad \langle \mathcal{J}'(x'), \varphi'(x') \rangle = \langle \Lambda \mathcal{J}(x), \varphi(x) \rangle$$

for any test function $\varphi(x) \in C_0^\infty(\mathbb{R}^4)$, where $\varphi'(x') := \varphi(x)$, $x' = \Lambda x$. Here we have used that $|\det \Lambda| = 1$ by (22.7). Since $\mathcal{J}(x)$ is a measure, it suffices to check (23.8) for characteristic functions $\varphi(x)$ of any open bounded set $\Omega \subset \mathbb{R}^4$. Then (23.8) becomes,

$$(23.9) \quad \int_{\Omega'} \mathcal{J}'(x') d^4 x' = \Lambda \int_{\Omega} \mathcal{J}(x) d^4 x.$$

Let us assume that the set $\{x \in \Omega : \mathbf{x} = \mathbf{x}(x_0)\}$ is an interval of the trajectory $\{(x_0, \mathbf{x}(x_0)) : a < x_0 < b\}$. Then integrating first in $d\mathbf{x}$, we get

$$(23.10) \quad \int_{\Omega} \mathcal{J}(x) d^4 x = e \int_a^b (1, \mathbf{v}(x_0)) dx_0 = e(x_0, \mathbf{x}(x_0)) \Big|_a^b.$$

It implies (23.9) since $(x_0, \mathbf{x}(x_0))$ is an invariant 4-vector, i.e. $(x'_0, \mathbf{x}'(x'_0)) := \Lambda(x_0, \mathbf{x}(x_0))$. ■

Remarks 23.4 *i) The charge invariance (23.6) implies the transformation (23.7) for the convective current. For all other currents, the transformation is postulated.*

*ii) The density $(1, \mathbf{v}(x_0))$ is **not** an invariant 4-vector. On the other hand, the differential form $(1, \mathbf{v}(x_0))dx_0$ is Lorentz-invariant which follows from the integration (23.10).*

Let us note that the form $(1, \mathbf{v}(x_0))dx_0$ admits the factorization

$$(23.11) \quad (1, \mathbf{v}(x_0))dx_0 = \frac{(1, \mathbf{v}(x_0))}{\sqrt{1 - \mathbf{v}^2(x_0)}} dx_0 \sqrt{1 - \mathbf{v}^2(x_0)},$$

where the fraction in RHS is an invariant 4-vector while the remaining factor is an invariant differential form. Indeed,

$$(23.12) \quad w(x_0) := \frac{(1, \mathbf{v}(x_0))}{\sqrt{1 - \mathbf{v}^2(x_0)}} = \frac{d(x_0, \mathbf{x}(x_0))}{dx_0 \sqrt{1 - \mathbf{v}^2(x_0)}} = \frac{d(x_0, \mathbf{x}(x_0))}{\sqrt{(dx_0)^2 - (d\mathbf{x}(x_0))^2}}$$

is an invariant 4-vector since the Lorentz interval $(dx_0)^2 - (d\mathbf{x}(x_0))^2$ is Lorentz-invariant. Respectively, $dx_0 \sqrt{1 - \mathbf{v}^2(x_0)} = \sqrt{(dx_0)^2 - (d\mathbf{x}(x_0))^2}$ is a Lorentz-invariant differential form.

Definition 23.5 *i) The vector $w(x_0)$ from (23.12) is the **4-dimensional velocity** of a particle.*

*ii) The function $\tau(b) := \int_a^b \sqrt{(dx_0)^2 - (d\mathbf{x}(x_0))^2} = \int_a^b \sqrt{1 - \mathbf{v}^2(x_0)} dx_0$ is the **proper time** of the particle along the trajectory.*

Now (23.11) can be written as

$$(23.13) \quad (1, \mathbf{v}(x_0))dx_0 = w(x_0)d\tau(x_0).$$

Corollary 23.6 *The formulas (23.3) imply the corresponding transformation for the Maxwell tensor (4.14):*

$$(23.14) \quad \mathcal{F}'^{\mu'\nu'}(x) = \Lambda^{\mu'}_{\mu} \Lambda^{\nu'}_{\nu} \mathcal{F}^{\mu\nu}(x), \quad x'_{\mu'} = \Lambda^{\mu}_{\mu'} x_{\mu}.$$

In the matrix form

$$(23.15) \quad \mathcal{F}'(x) = \Lambda \mathcal{F}(x) \Lambda^t.$$

Example 23.7 Applying this formula to the Lorentz boost (21.22), we get the Lorentz transformation for the Maxwell fields :

$$(23.16) \quad \left\{ \begin{array}{ll} \mathbf{E}'_1(x') = \mathbf{E}_1(x), & \mathbf{B}'_1(x') = \mathbf{B}_1(x), \\ \mathbf{E}'_2(x') = \frac{\mathbf{E}_2(x) - \beta \mathbf{B}_3(x)}{\sqrt{1 - \beta^2}}, & \mathbf{B}'_2(x') = \frac{\mathbf{B}_2(x) + \beta \mathbf{E}_3(x)}{\sqrt{1 - \beta^2}}, \\ \mathbf{E}'_3(x') = \frac{\mathbf{E}_3(x) + \beta \mathbf{B}_2(x)}{\sqrt{1 - \beta^2}}, & \mathbf{B}'_3(x') = \frac{\mathbf{B}_3(x) - \beta \mathbf{E}_2(x)}{\sqrt{1 - \beta^2}}. \end{array} \right.$$

Equivalently,

$$(23.17) \quad \mathbf{E}'(x') = \frac{\mathbf{E}(x) + \frac{\mathbf{v}}{c} \times \mathbf{B}(x)}{\sqrt{1 - \beta^2}}, \quad \mathbf{B}'(x') = \frac{\mathbf{B}(x) - \frac{\mathbf{v}}{c} \times \mathbf{E}(x)}{\sqrt{1 - \beta^2}}.$$

Exercise 23.8 Check the formulas (23.16). **Hints:** The formulas follow from (4.15) by (23.15) with the Lorentz matrix (21.22).

Part VI

Relativistic Dirac Theory

24 Relativistic Equation for Electron Field

The Special Relativity of Einstein has to be extended to the quantum mechanical theory, as well. We introduce the Dirac equation which is the relativistic covariant generalization of the Schrödinger equation.

The Schrödinger equation obviously is not invariant with respect to the Lorentz group. One possible choice is the Klein-Gordon equation

$$(24.1) \quad \frac{\hbar^2}{c^2} \partial_t^2 \psi(x) = \hbar^2 \Delta \psi(x) - \mu^2 c^2 \psi(x), \quad x \in \mathbb{R}^4,$$

which is Lorentz-invariant as well as the wave equation (21.23). However, it leads to negative energies which is not satisfactory from a physical point of view. Namely, the Klein-Gordon equation is obtained from the relativistic energy-momentum relation

$$(24.2) \quad \frac{E^2}{c^2} = \mathbf{p}^2 + \mu^2 c^2$$

by the Schrödinger identification $E \mapsto i\hbar \partial_t$, $\mathbf{p} \mapsto -i\hbar \nabla_{\mathbf{x}}$. In the Fourier transform, $\hat{\psi}(x^0, \mathbf{p}) := \int e^{i\frac{\mathbf{p}\mathbf{x}}{\hbar}} \psi(x_0, \mathbf{x}) d\mathbf{x}$, the Klein-Gordon equation becomes

$$(24.3) \quad \hbar^2 \frac{\partial^2}{(\partial x^0)^2} \hat{\psi}(x_0, \mathbf{p}) = -\mathbf{p}^2 \hat{\psi}(x_0, \mathbf{p}) - \mu^2 c^2 \hat{\psi}(x_0, \mathbf{p}), \quad x_0 \in \mathbb{R}, \mathbf{p} \in \mathbb{R}^3.$$

Then the solutions are linear combinations of $e^{i\frac{E}{\hbar}t}$ where $E/c = \pm\sqrt{\mathbf{p}^2 + \mu^2 c^2}$ coincides with (24.2). The solutions with $E/c = -\sqrt{\mathbf{p}^2 + \mu^2 c^2}$ seem to correspond to negative energies unbounded from below, hence the physical interpretation of the Klein-Gordon equation requires an additional analysis.

This is why Dirac tried to find a relativistic invariant equations of the first order in time, like the Schrödinger equation, to avoid the negative roots. Let us follow Dirac's arguments. First, the relativistic invariance requires then the first order in space. Second, we will look for a linear equation of the form

$$(24.4) \quad \gamma^\alpha i\hbar \frac{\partial}{\partial x^\alpha} \psi(x) = \mu c \psi(x), \quad x \in \mathbb{R}^4.$$

Then the corresponding energy-momentum relation can be written as

$$(24.5) \quad \gamma(p) = \mu c, \quad p \in \mathbb{R}^4,$$

where $p_\alpha := (E/c, -\mathbf{p})$ and $\gamma(p)$ is a linear form:

$$(24.6) \quad \gamma(p) = \gamma^\alpha p_\alpha, \quad p \in \mathbb{R}^4.$$

The third condition is a "correspondence principle": the equation (24.4) must imply the Klein-Gordon equation (24.1). Namely, applying the operator $\gamma^\alpha i\hbar \frac{\partial}{\partial x^\alpha}$ to both sides, we get

$$(24.7) \quad [\gamma^\alpha i\hbar \frac{\partial}{\partial x^\alpha}]^2 \psi(x) = \mu^2 c^2 \psi(x), \quad x \in \mathbb{R}^4.$$

Then the correspondence principle is equivalent to the algebraic identity

$$(24.8) \quad \gamma^2(p) = p_0^2 - \mathbf{p}^2, \quad p = (p_0, \mathbf{p}) \in \mathbb{R}^4$$

since (24.2) can be written as $p_0^2 - \mathbf{p}^2 = \mu^2 c^2$. Dirac's extra idea was the choice of the coefficients γ^α in the matrix algebra since the scalar coefficients do not exist. The existence of the scalar coefficients would mean that the polynomial $p_0^2 - \mathbf{p}^2$ is reducible which is not true.

Exercise 24.1 Check that (24.8) is impossible with any scalar coefficients γ^α .

Theorem 24.2 In 2×2 block form, the matrices

$$(24.9) \quad \gamma(p) = \begin{pmatrix} p_0 & \boldsymbol{\sigma} \cdot \mathbf{p} \\ -\boldsymbol{\sigma} \cdot \mathbf{p} & -p_0 \end{pmatrix}$$

satisfy the identity (24.8), where $\boldsymbol{\sigma} := (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli spin matrices.

Proof By direct multiplication of 2×2 block matrices, we get

$$(24.10) \quad \gamma^2(p) = \begin{pmatrix} p_0^2 - (\boldsymbol{\sigma} \cdot \mathbf{p})^2 & 0 \\ 0 & p_0^2 - (\boldsymbol{\sigma} \cdot \mathbf{p})^2 \end{pmatrix}.$$

It remains to use that $(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = |\mathbf{p}|^2$. ■

Now let us consider the matrices $\gamma^\alpha = \gamma(e_\alpha)$, where $e_0 = (1, 0, 0, 0)$, etc. From (24.9), we get

$$(24.11) \quad \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}, \quad j = 1, 2, 3.$$

Proposition 24.3 The matrices γ^α satisfy

$$(24.12) \quad \begin{cases} (\gamma^0)^2 = 1, & (\gamma^j)^2 = -1, \quad j = 1, 2, 3, \\ \gamma^\alpha \gamma^\beta + \gamma^\beta \gamma^\alpha = 0, & \alpha \neq \beta. \end{cases}$$

Proof Let us rewrite (24.8) in the form

$$(24.13) \quad \gamma^2(p) = g(p), \quad p \in \mathbb{R}^4,$$

where $g(p) := p_0^2 - \mathbf{p}^2$. It implies

$$(24.14) \quad \gamma(p)\gamma(q) + \gamma(q)\gamma(p) = 2g(p, q),$$

where $g(p, q) = p_0q_0 - \mathbf{p}\mathbf{q}$ is the corresponding symmetric bilinear form. In particular, for $p = e_\alpha$ and $q = e_\beta$

$$(24.15) \quad \gamma^\alpha \gamma^\beta + \gamma^\beta \gamma^\alpha = 2g(e_\alpha, e_\beta)$$

which implies (24.12). ■

Remark 24.4 Obviously the matrices (24.11) are not unique solutions to the relations (24.12): for example, we can replace γ_α by $-\gamma_\alpha$ for certain indexes α . We will prove below the Pauli Theorem: the matrices γ_α are unique up to a change of an orthonormal basis.

Let us rewrite the Dirac equation (24.4) in a ‘‘Schrödinger form’’. First, it is equivalent to

$$(24.16) \quad i\hbar\gamma^0\dot{\psi}(t, \mathbf{x}) = c(\mu c - i\hbar\gamma^j\frac{\partial}{\partial x^j})\psi(t, \mathbf{x}).$$

Second, multiplying by γ^0 and using that $(\gamma^0)^2 = 1$, we get

$$(24.17) \quad i\hbar\dot{\psi}(t, \mathbf{x}) = \mathcal{H}_D\psi(t, \mathbf{x}) := c\gamma^0(\mu c - i\hbar\gamma^j\frac{\partial}{\partial x^j})\psi(t, \mathbf{x}),$$

where the operator \mathcal{H}_D is called the *free Dirac Hamiltonian*.

25 Problem of Negative Energies for Dirac Equation

We prove the energy conservation for the Dirac equation and check that the energy is not bounded from below. The Dirac equation (24.17) is a Hamiltonian system with the Hamilton functional

$$(25.1) \quad H(\psi) := \langle \psi(\mathbf{x}), \mathcal{H}_D \psi(\mathbf{x}) \rangle,$$

in analogy with the Schrödinger equation. It is conserved, i.e.

$$(25.2) \quad H(\psi(t, \cdot)) = \text{const}, \quad t \in \mathbb{R}$$

for the solutions to (24.17).

Exercise 25.1 Check (25.2). **Hint:** Differentiate (25.1) and use the Dirac equation (24.17) and the symmetry of the Dirac operator \mathcal{H}_D .

The quadratic form (25.1) is not positive definite, hence energy conservation (25.2) does not provide an a priori estimate for the solutions. To see this, it is useful to split each Dirac spinor into a pair of two-component vectors

$$(25.3) \quad \psi(\mathbf{x}) = \begin{pmatrix} \psi_+(\mathbf{x}) \\ \psi_-(\mathbf{x}) \end{pmatrix}.$$

Define the Fourier transform

$$(25.4) \quad \hat{\psi}(\mathbf{p}) = \int e^{i\mathbf{p}\mathbf{x}/\hbar} \psi(\mathbf{x}) d\mathbf{x}, \quad \mathbf{p} \in \mathbb{R}^3.$$

Then the quadratic form (25.1) becomes by the Parseval identity

$$\begin{aligned} H(\psi) &= (2\pi)^{-3} \langle \hat{\psi}(\mathbf{p}), c\gamma^0 \begin{pmatrix} \mu c & -\sigma \cdot \mathbf{p} \\ \sigma \cdot \mathbf{p} & \mu c \end{pmatrix} \hat{\psi}(\mathbf{p}) \rangle = c(2\pi)^{-3} \langle \hat{\psi}(\mathbf{p}), \begin{pmatrix} \mu c & -\sigma \cdot \mathbf{p} \\ -\sigma \cdot \mathbf{p} & -\mu c \end{pmatrix} \hat{\psi}(\mathbf{p}) \rangle \\ &= c(2\pi)^{-3} [\mu c \langle \hat{\psi}_+(\mathbf{p}), \hat{\psi}_+(\mathbf{p}) \rangle - 2 \langle \hat{\psi}_+(\mathbf{p}), \sigma \cdot \mathbf{p} \hat{\psi}_-(\mathbf{p}) \rangle - \mu c \langle \hat{\psi}_-(\mathbf{p}), \hat{\psi}_-(\mathbf{p}) \rangle]. \end{aligned}$$

In particular,

$$(25.5) \quad H \begin{pmatrix} \psi_+(\mathbf{x}) \\ 0 \end{pmatrix} = (2\pi)^{-3} \mu c^2 \langle \hat{\psi}_+(\mathbf{p}), \hat{\psi}_+(\mathbf{p}) \rangle, \quad H \begin{pmatrix} 0 \\ \psi_-(\mathbf{x}) \end{pmatrix} = -(2\pi)^{-3} \mu c^2 \langle \hat{\psi}_-(\mathbf{p}), \hat{\psi}_-(\mathbf{p}) \rangle$$

The negative energy might lead to an instability of the Dirac dynamics due to a possible transition of the solution to the states $\begin{pmatrix} 0 \\ \psi_-(\mathbf{x}) \end{pmatrix}$. On the other hand, this instability has never been proved.

Dirac suggested that the transition of all particles is forbidden by the Pauli exclusion principle since almost all states with negative energy have been occupied long ago. (A more satisfactory solution of the problem of negative energies, which is applicable also for bosons, is provided by Quantum Field Theory.) On the other hand, by the Dirac theory, the transitions for certain particles are possible, and the 'negative states' can be interpreted as the states with positive energy for *antiparticles* which are *positrons*, i.e., 'electrons with positive charge' $-e$. The positrons have been discovered in *cosmic rays* by Anderson in 1932.

26 Angular Momentum for Dirac Equation

We prove the angular momentum conservation for the Dirac equation. The conservation justifies the Goudsmith-Uhlenbeck conjecture on the electron spin as an intrinsic property of the dynamical equations.

The conserved orbital momentum for the Schrödinger equation is defined by (6.17) and (6.18): $\mathbf{L} = \mathbf{L}(\psi) := \langle \psi, \hat{\mathbf{L}}\psi \rangle$, where $\hat{\mathbf{L}} = -i\hbar\mathbf{x} \times \nabla$. For the solutions to the Dirac equation the orbital momentum generally is not conserved, since the operator $\hat{\mathbf{L}}$ does not commute with the Dirac operator H_D . Hence, for the Dirac equation the definition of the angular momentum requires a modification.

Definition 26.1 For the Dirac equation the angular momentum is $\mathbf{J} = \mathbf{J}(\psi) = \langle \psi, \hat{\mathbf{J}}\psi \rangle$ where

$$(26.1) \quad \hat{\mathbf{J}} := \hat{\mathbf{L}} + \frac{1}{2}\hbar\tilde{\sigma}, \quad \tilde{\sigma} := \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}.$$

Theorem 26.2 The angular momentum \mathbf{J} is conserved for the solutions to the Dirac equation,

$$(26.2) \quad \mathbf{J}(\psi(t, \cdot)) = \text{const}, \quad t \in \mathbb{R}.$$

Proof Differentiating, we obtain similarly to (6.26),

$$(26.3) \quad \begin{aligned} \frac{d}{dt}\mathbf{J}(\psi(t)) &= \langle \dot{\psi}(t), \hat{\mathbf{J}}\psi(t) \rangle + \langle \psi(t), \hat{\mathbf{J}}\dot{\psi}(t) \rangle = -\langle \frac{i}{\hbar}\mathcal{H}_D\psi(t), \hat{\mathbf{J}}\psi(t) \rangle - \langle \psi(t), \hat{\mathbf{J}}\frac{i}{\hbar}\mathcal{H}_D\psi(t) \rangle \\ &= -\frac{i}{\hbar} \left[\langle \mathcal{H}_D\psi(t), \hat{\mathbf{J}}\psi(t) \rangle - \langle \psi(t), \hat{\mathbf{J}}\mathcal{H}_D\psi(t) \rangle \right] = -\frac{i}{\hbar} \langle \psi(t), [\mathcal{H}_D, \hat{\mathbf{J}}]\psi(t) \rangle. \end{aligned}$$

It suffices to check the commutation $[\mathcal{H}_D, \hat{\mathbf{J}}] = 0$. First, we know the commutators $[\hat{\mathbf{L}}_k, p_j] = i\hbar\epsilon_{kjl}p_l$, where $p_j := i\hbar\partial_j$ and ϵ_{kjl} is the totally antisymmetric tensor. Therefore,

$$(26.4) \quad [\hat{\mathbf{L}}_k, \mathcal{H}_D] = -c\hbar\gamma^0\gamma^j[\mathbf{L}_k, p_j] = -ic\hbar\gamma^0\gamma^j\epsilon_{kjl}p_l.$$

This shows that \mathcal{H}_D does not commute with the orbital angular momentum operators $\hat{\mathbf{L}}_k$, hence the orbital momentum \mathbf{L} generally is not conserved. It remains to calculate the commutators $[\tilde{\sigma}, \mathcal{H}_D]$. First we note that

$$\gamma^0\gamma^l = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & \sigma_l \\ -\sigma_l & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_l \\ \sigma_l & 0 \end{pmatrix}.$$

Hence,

$$[\tilde{\sigma}_k, \mathcal{H}_D] = [\tilde{\sigma}_k, \mu c^2\gamma^0 - c\gamma^0\gamma^l p_l] = -c \begin{pmatrix} 0 & [\sigma_k, \sigma_l] \\ [\sigma_k, \sigma_l] & 0 \end{pmatrix} p_l.$$

The commutation relations for the Pauli spin matrices allows us to reduce this to

$$-2ic\epsilon_{klj} \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} p_l = -2ic\epsilon_{klj}\gamma^0\gamma^j p_l.$$

Multiplying this by $\hbar/2$ and adding (26.4), we get the commutation $[\mathcal{H}_D, \hat{\mathbf{J}}] = 0$ by the antisymmetry of ϵ_{klj} . ■

Remarks 26.3 *i) Theorem 26.2 means that $\frac{1}{2}\hbar\tilde{\sigma}$ represents an intrinsic spinor angular momentum of the relativistic electron.*

ii) We will demonstrate below that the coupling of the Dirac equation to the magnetic field provides automatically the correct Landé factor $g = 2$ for the spinor angular momentum.

27 Pauli Theorem

We prove the Pauli theorem on the uniqueness of the Dirac matrices. The theorem states that the Dirac matrices (24.11) are the unique representations of the relations (24.12), up to equivalence.

Theorem 27.1 *Let γ^α , $\alpha = 0, \dots, 3$ be operators on a finite-dimensional complex vector space V , $\dim V \leq 4$, which satisfy the relations (24.12). Then $\dim V = 4$ and the operators γ^α , for a suitable choice of a basis, take the matrix form (24.11).*

Proof *Step i)* A key idea of the proof is the following simple characterization of the basis vectors. Namely, for the standard Dirac matrices (24.11), the matrix $\gamma^1\gamma^2$ is diagonal and hence commutes with γ^0 which is also diagonal: γ^0 and $\gamma^1\gamma^2$ have the diagonal block matrix form

$$(27.1) \quad \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^1\gamma^2 = \begin{pmatrix} -i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix}.$$

Therefore, the basis vectors e_0, \dots, e_3 are common eigenvectors of the matrices γ^0 and $\gamma^1\gamma^2$ with the eigenvalues 1 and $-i$, 1 and i , -1 and $-i$, -1 and i respectively.

Step ii) Now apply this observation to the general matrices γ^α from the theorem. The anticommutation relations from (24.12) imply that the matrices γ^0 and $\gamma^1\gamma^2$ commute with each other:

$$(27.2) \quad \gamma^0\gamma^1\gamma^2 = \gamma^1\gamma^2\gamma^0.$$

Exercise 27.2 *Check this commutation. Solution:* $\gamma^0\gamma^1\gamma^2 = -\gamma^1\gamma^0\gamma^2 = \gamma^1\gamma^2\gamma^0$.

Hence there exists at least one common eigenvector Ω_1 for both (since V is a complex vector space!):

$$(27.3) \quad \gamma^0\Omega_1 = \alpha\Omega_1 \quad \text{and} \quad \gamma^1\gamma^2\Omega_1 = \beta\Omega_1$$

where α and β are suitable complex numbers.

Step iii) $\alpha^2 = 1$ since $(\gamma^0)^2 = 1$, and similarly $\beta^2 = -1$ since $(\gamma^1\gamma^2)^2 = -(\gamma^1)^2(\gamma^2)^2 = -1$. Hence, $\alpha = \pm 1$ and $\beta = \pm i$. Let us check that all four combinations of the signs are possible for suitable eigenvectors Ω_1 . Namely,

$$\gamma^0\gamma^3\Omega_1 = -\gamma^3\gamma^0\Omega_1 = -\alpha\gamma^3\Omega_1,$$

and

$$\gamma^1\gamma^2\gamma^3\Omega_1 = \gamma^3\gamma^1\gamma^2\Omega_1 = \beta\gamma^3\Omega_1,$$

hence the vector $\Omega_3 := \gamma^3\Omega_1$ is also a common eigenvector with the eigenvalues $-\alpha$ and β . Similarly, $\Omega_2 := -\gamma^3\gamma^1\Omega_1$ resp. $\Omega_4 := -\gamma^1\Omega_1$ are common eigenvectors of the operators γ^0 and $\gamma^1\gamma^2$ with the eigenvalues α and $-\beta$ resp. $-\alpha$ and $-\beta$. Since all four possible signs occur, we may permute the four vectors to ensure that $\alpha = 1$ and $\beta = -i$.

Step iv) On the subspace with basis $\Omega_1, \Omega_2, \Omega_3$ and Ω_4 the operators γ^0 and $\gamma^1\gamma^2$ have the diagonal block matrix form (27.1). Moreover, in this basis the operators γ^1 and γ^3 have the form

$$(27.4) \quad \gamma^1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}$$

which coincide with (24.11). From these we may check that $\gamma^2 = -\gamma^1(\gamma^1\gamma^2)$ also has the desired form.

Step v) The vectors Ω_j , $j = 1, \dots, 4$, span the space V since $\dim V \leq 4$. ■

Exercise 27.3 Check the formulae (27.1) for γ^0 and $\gamma^1\gamma^2$.

Exercise 27.4 Check the formulae (27.4) for γ^1 and γ^3 .

Corollary 27.5 For any Lorentz transformation Λ , there exists a nondegenerate matrix $\Gamma(\Lambda) \in GL(4, \mathbb{C})$ such that

$$(27.5) \quad \gamma(\Lambda p) = \Gamma(\Lambda)\gamma(p)\Gamma^{-1}(\Lambda), \quad p \in \mathbb{R}^4.$$

Proof (24.13) implies

$$(27.6) \quad \gamma^2(\Lambda p) = g(\Lambda p) = g(p), \quad p \in \mathbb{R}^4,$$

since Λ is a Lorentz transformation. Hence, the matrices $\gamma(\Lambda e_\alpha)$ satisfy the relations (24.12) as well as $\gamma^\alpha := \gamma(e_\alpha)$. Therefore, by the Pauli Theorem, we have

$$(27.7) \quad \gamma(\Lambda e_\alpha) = \Gamma(\Lambda)\gamma(e_\alpha)\Gamma^{-1}(\Lambda), \quad \alpha = 0, \dots, 3,$$

where $\Gamma(\Lambda)$ is an invertible operator in \mathbb{R}^4 (which transforms the vector e_α into Ω_α , $\alpha = 0, \dots, 3$). Then (27.5) follows by linearity. ■

28 Lorentz Covariance

We justify the Einstein postulate of Special Relativity Theory for the Dirac equation: the equation takes an identical form in all Inertial Frames.

Let us consider two frames of reference of two observers related by a Lorentz transformation: $x' = \Lambda x$. By a natural extension of the Einstein postulate **E**, the Dirac equation (24.4) has the same form in both frames of reference. This extended postulate allows us to determine the corresponding transformation of the wave function. The next theorem gives a transformation of the wave function which leaves the Dirac equation invariant.

Theorem 28.1 *Let $\psi(x)$ be a solution to the Dirac equation (24.4) and*

$$(28.1) \quad \psi'(x') := \Gamma(\Lambda^\#)\psi(x), \quad x \in \mathbb{R}^4,$$

where $x' = \Lambda x$ and $\Lambda^\# := (\Lambda^t)^{-1}$ where Λ^t is the transposed matrix to Λ . Then the function $\psi'(x')$ is also a solution to the Dirac equation.

Proof Let us translate the Dirac equation (24.4) into the Fourier transform

$$(28.2) \quad \hat{\psi}(p) := \int e^{\frac{ipx}{\hbar}} \psi(x) dx, \quad p \in \mathbb{R}^4,$$

where $px := p_\alpha x^\alpha$. Then (24.4) becomes

$$(28.3) \quad \gamma(p)\hat{\psi}(p) = \mu c \hat{\psi}(p), \quad p \in \mathbb{R}^4.$$

The Fourier transform translates (28.1) into

$$(28.4) \quad \hat{\psi}'(p') = \Gamma(\Lambda^\#)\hat{\psi}(p), \quad p \in \mathbb{R}^4,$$

where $p = \Lambda^t p'$.

Exercise 28.2 *Check (28.4). Hint: formally, we have*

$$(28.5) \quad \hat{\psi}'(p') := \int e^{\frac{ip'x'}{\hbar}} \psi'(x') dx' = \int e^{\frac{ip'\Lambda x}{\hbar}} \Gamma(\Lambda^\#)\psi(x) |\det \Lambda| dx = \Gamma(\Lambda^\#) \int e^{\frac{i\Lambda^t p' x}{\hbar}} \psi(x) dx = \Gamma(\Lambda^\#)\hat{\psi}(p)$$

since $|\det \Lambda| = 1$ for the Lorentz transformation Λ .

Now express (28.3) in terms of the wave function $\hat{\psi}'(p')$:

$$(28.6) \quad \gamma(p)\Gamma^{-1}(\Lambda^\#)\hat{\psi}'(p') = \mu c \Gamma^{-1}(\Lambda^\#)\hat{\psi}'(p'), \quad p' \in \mathbb{R}^4.$$

This is equivalent to the Dirac equation (28.3) iff

$$(28.7) \quad \Gamma(\Lambda^\#)\gamma(p)\Gamma^{-1}(\Lambda^\#) = \gamma(p'), \quad p' \in \mathbb{R}^4.$$

This is equivalent to (27.5) for $\Lambda^\#$ instead of Λ since $p' = \Lambda^\# p$. Finally, it is true since $\Lambda^\#$ also belongs to the Lorentz group. ■

Exercise 28.3 *Check that $\Lambda^\#$ is a Lorentz transformation for any $\Lambda \in L$.*

Remarks 28.4 i) The formal calculations (28.5) can be justified by the properties of the Fourier transformation of tempered distributions. This is necessary since the integrals in (28.5) generally never converge for the solutions to the Dirac equation by energy and charge conservations (see below).

ii) The theorem implies that the "variance" rule (28.1) leaves the Dirac equation unchanged. This means, by definition, that the Dirac equation is *covariant* with respect to the Lorentz group.

29 Lorentz Transformation of Spinors

We give a construction of the operator $\Gamma(\Lambda)$ acting on the spinor wave functions. For the rotations the operator will be calculated explicitly.

29.1 Factorization of Lorentz Transformations

First let us derive a useful formula for the transformations of the Dirac matrices.

Lemma 29.1 *i) For any two vectors $p, q \in \mathbb{R}^4$ with q non-null w.r.t. g (i.e. $g(q, q) \neq 0$), we have*

$$(29.1) \quad \gamma(q)\gamma(p)\gamma^{-1}(q) = \gamma(R_q p),$$

where R_q is a 'reflection' $R_q p = 2 \frac{g(q, p)}{g(q, q)} q - p$.

ii) The transformation R_q is in the Lorentz group but not proper, i.e. $\det R_q = -1$.

Proof i) Since $\gamma^2(q) = g(q, q)$, we have $\gamma^{-1}(q) = \gamma(q)/g(q, q)$. Therefore, multiplying the relation

$$\gamma(q)\gamma(p) + \gamma(p)\gamma(q) = 2g(q, p)$$

on the right by $\gamma^{-1}(q)$, we obtain

$$\gamma(q)\gamma(p)\gamma^{-1}(q) + \gamma(p) = 2 \frac{g(q, p)}{g(q, q)} \gamma(q).$$

This implies (29.1) by the linearity of γ .

ii) It is easy to check that R_q is in the Lorentz group, i.e.

$$(29.2) \quad g(R_q p, R_q p) = g(p, p), \quad p \in \mathbb{R}^4.$$

It remains to check that

$$(29.3) \quad \det R_q = -1.$$

This is obvious since

i) R_q preserves the components along q , i.e. $R_q q = q$.

ii) R_q reverses the components g -orthogonal to q , i.e. $R_q p = -p$ if $g(q, p) = 0$. ■

Exercise 29.2 Check (29.2).

Example 29.3 $q = (\pm 1, 0)$: then $R_q p = (2p_0, 0, 0, 0) - p = (p_0, -\mathbf{p})$ for $p = (p^0, \mathbf{p})$.

Example 29.4 $q = (0, \mathbf{q})$ with \mathbf{q} a unit vector: then $g(q, q) = -|\mathbf{q}|^2 = -1$ and $R_q p = 2(\mathbf{q} \cdot \mathbf{p})q - p = (-p_0, 2(\mathbf{q} \cdot \mathbf{p})\mathbf{q} - \mathbf{p})$. Hence the spatial component along \mathbf{q} is unchanged while the orthogonal components are reversed, which is precisely the effect of a rotation through π about the axis \mathbf{q} .

Now let us factorize the Lorentz transformations in two reflections.

Lemma 29.5 *Let Λ be a proper Lorentz transformation that fixes two linearly independent vectors v_1 and v_2 . Then it admits the factorization*

$$(29.4) \quad \Lambda v = R_w R_u v, \quad v \in \mathbb{R}^4.$$

Here u is a non-null vector (i.e. $g(u, u) \neq 0$), g -orthogonal to v_1 and v_2 , and we set $w = \Lambda u + u$ if $\Lambda u \neq -u$. Otherwise, we take w which is g -orthogonal to all three vectors v_1, v_2 and u .

Proof First let us note that u can be chosen non-null. Indeed, the plane Π , g -orthogonal to v_1 and v_2 , is two-dimensional. Hence, its intersection with the null cone is at most one-dimensional cone since the null cone does not contain two-dimensional planes. Therefore, all vectors $u \in \Pi$ are non-null except for at most two lines.

Second, note that

$$(29.5) \quad g(\Lambda u, v_j) = g(\Lambda u, \Lambda v_j) = g(u, v_j) = 0.$$

since Λ fixes v_j and preserves g . Hence, Λu and $w = \Lambda u + u$ are g -orthogonal to v_1 and v_2 .

Further consider the case $\Lambda u \neq -u$. Then $\Lambda + 1$ is an invertible operator. Namely, Λ has two eigenvectors v_1, v_2 with the eigenvalue 1. If it has an eigenvector v_3 with the eigenvalue -1 , then also a vector v_4 g -orthogonal to v_1, v_2 and v_3 is an eigenvector with the eigenvalue -1 . Then also $\Lambda u = -u$ which contradicts our assumption.

Therefore, $(\Lambda + 1)\Pi$ is also a two-dimensional plane. Hence, by the same arguments, $w = \Lambda u + u$ is a non-null vector for almost all $u \in \Pi$.

Now let us check the identity (29.4). First, the composition $R_w R_u$ fixes v_1 and v_2 since each factor reverses both vectors. Hence, (29.4) holds for $v = v_1$ and $v = v_2$. Further, let us prove (29.4) for $v = u$. Namely,

$$(29.6) \quad \begin{aligned} g(w, w) &= g(\Lambda u, \Lambda u) + 2g(\Lambda u, u) + g(u, u) \\ &= g(u, u) + 2g(\Lambda u, u) + g(u, u) \\ &= 2[g(u, u) + g(\Lambda u, u)] = 2g(u, w). \end{aligned}$$

Therefore,

$$(29.7) \quad R_w u = \frac{2g(u, w)}{g(w, w)} w - u = w - u = \Lambda u.$$

Then also $\Lambda u = R_w R_u u$ since $R_u u = u$.

So (29.4) holds for three linearly independent vectors v_1, v_2 and u . Hence their action on the whole space \mathbb{R}^4 is identical since both $R_w R_u$ and Λ are proper Lorentz transformations. The case $\Lambda u = -u$ can be checked similarly. ■

Exercise 29.6 Check the lemma for the case $\Lambda u = -u$.

The next lemma demonstrates that each Lorentz transformation admits a factorization into three ones, each preserving two linearly independent vectors. Let us recall that the rotation group $SO(3)$ is a natural subgroup of the Lorentz group L : for each $R \in SO(3)$ the corresponding Lorentz transformation is given by (22.12).

Lemma 29.7 Let Λ be a proper Lorentz transformation. Then it admits the factorization

$$(29.8) \quad \Lambda = \hat{R}_1 B \hat{R}_2,$$

where $R_1, R_2 \in SO(3)$ and B is a proper boost of type (22.15).

Proof Let us consider the vector $\Lambda e_0 = (v_0, \mathbf{v})$, where $e_0 := (1, 0, 0, 0)$. There exists a rotation $Q_1 \in SO(3)$ such that $Q_1 \mathbf{v} = (v_1, 0, 0)$. Then the vector $e'_0 := \hat{Q}_1 \Lambda e_0$ has the form $(v_0, v_1, 0, 0)$ with $g(e'_0) = v_0^2 - v_1^2 = 1$ since $g(e_0) = 1$ and $\hat{Q}_1 \Lambda \in L$.

Exercise 29.8 Check that there exists a proper boost D of type (22.15) such that $De'_0 = e_0$. **Hint:** Choose D in the form Λ_+^+ if $v_0 > 0$ and Λ_-^- if $v_0 < 0$.

Now $D\hat{Q}_1\Lambda e_0 = e_0$, hence the Lorentz transformation $B\hat{R}_2\Lambda$ has a matrix of type

$$(29.9) \quad D\hat{Q}_1\Lambda = \begin{pmatrix} 1 & \mathbf{w} \\ 0 & Q_2 \end{pmatrix},$$

Exercise 29.9 Check that $\mathbf{w} = 0$. **Hint:** use (22.6) and (22.11) for the matrix $D\hat{Q}_1\Lambda$.

Therefore, $Q_2 \in SO(3)$ and (29.9) implies that $D\hat{Q}_1\Lambda = \hat{Q}_2$, hence

$$(29.10) \quad \Lambda = \hat{Q}_2 D^{-1} \hat{Q}_1^{-1}. \quad \blacksquare$$

Lemmas 29.5 and 29.7 imply that each proper Lorentz transformation Λ admits a factorization

$$\Lambda = R_{q_1} \dots R_{q_6}.$$

Corollary 29.10 The matrix $\Gamma(\Lambda) := \gamma(q_1) \dots \gamma(q_6)$ satisfies

$$(29.11) \quad \gamma(\Lambda p) = \Gamma(\Lambda) \gamma(p) \Gamma^{-1}(\Lambda), \quad p \in \mathbb{R}^4.$$

Proof Lemma 29.1 implies by induction that

$$(29.12) \quad \Gamma(\Lambda) \gamma(p) \Gamma^{-1}(\Lambda) = \gamma(q_1) \dots \gamma(q_6) \gamma(p) \gamma^{-1}(q_6) \dots \gamma^{-1}(q_1) = \gamma(R_{q_1} \dots R_{q_6} p) = \gamma(\Lambda p). \quad \blacksquare$$

29.2 Rotations of Dirac Spinors

This corollary allows us to construct explicitly the matrix $\Gamma(\Lambda)$. Consider, for example, a rotation $\Lambda = \hat{R}(\theta \mathbf{n})$ through $\theta \in (-\pi, \pi)$ about an axis \mathbf{n} . Then (28.1) becomes

$$(29.13) \quad \psi'(x') := \Gamma(\hat{R}(\theta \mathbf{n})) \psi(\hat{R}(-\theta \mathbf{n})x'), \quad x' \in \mathbb{R}^4$$

since $\hat{R}^\#(\theta \mathbf{n}) = \hat{R}(\theta \mathbf{n})$ for the orthogonal matrix $R(\theta \mathbf{n})$. The rotation $\hat{R}(\theta \mathbf{n})$ fixes two vectors \mathbf{n} and e_0 . Hence we can apply Lemma 29.5 to factorize $\hat{R}(\theta \mathbf{n})$ and then Corollary 29.10 to construct $\Gamma(\hat{R}(\theta \mathbf{n}))$.

According to Lemma 29.5, we choose any non-null vector \mathbf{u} which is perpendicular to \mathbf{n} and note that

$$(29.14) \quad \hat{R}(\theta \mathbf{n})\mathbf{u} = \cos \theta \mathbf{u} - \sin \theta \mathbf{n} \times \mathbf{u}$$

for the rotation about \mathbf{n} in a positive direction. According to Lemma 29.5, we define

$$(29.15) \quad \begin{aligned} \mathbf{w} = \mathbf{u} + \hat{R}(\theta \mathbf{n})\mathbf{u} &= (1 + \cos \theta)\mathbf{u} - \sin \theta \mathbf{n} \times \mathbf{u} \\ &= 2 \cos \frac{\theta}{2} \left[\cos \frac{\theta}{2} \mathbf{u} - \sin \frac{\theta}{2} \mathbf{n} \times \mathbf{u} \right]. \end{aligned}$$

It is a nonzero vector since $\theta \in (-\pi, \pi)$, hence $\hat{R}(\theta \mathbf{n}) = \hat{R}_{\mathbf{w}} \hat{R}_{\mathbf{u}}$ by Lemma 29.5. Therefore, by Corollary 29.10, we can choose

$$(29.16) \quad \begin{aligned} \Gamma(\hat{R}(\theta \mathbf{n})) = \gamma(\mathbf{w})\gamma(\mathbf{u}) &= \begin{pmatrix} 0 & \sigma \mathbf{w} \\ -\sigma \mathbf{w} & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma \mathbf{u} \\ -\sigma \mathbf{u} & 0 \end{pmatrix} \\ &= \begin{pmatrix} -(\sigma \mathbf{w})(\sigma \mathbf{u}) & 0 \\ 0 & -(\sigma \mathbf{w})(\sigma \mathbf{u}) \end{pmatrix}. \end{aligned}$$

Since the reflection $\hat{R}_{\mathbf{w}}$ is unaffected by any normalization of \mathbf{w} , let us redefine \mathbf{w} as the unit vector $-\cos \frac{\theta}{2} \mathbf{u} + \sin \frac{\theta}{2} \mathbf{n} \times \mathbf{u}$. Then

$$(29.17) \quad (\sigma \mathbf{w})(\sigma \mathbf{u}) = \mathbf{u} \mathbf{w} + i \sigma (\mathbf{w} \times \mathbf{u}) = -\cos \frac{\theta}{2} - i \sin \frac{\theta}{2} \sigma \mathbf{n} = -\exp\left(\frac{i\theta}{2} \sigma \mathbf{n}\right)$$

by the Euler trick. Finally,

$$(29.18) \quad \Gamma(\hat{R}(\theta\mathbf{n})) = \begin{pmatrix} \exp(\frac{i\theta}{2}\sigma\mathbf{n}) & 0 \\ 0 & \exp(\frac{i\theta}{2}\sigma\mathbf{n}) \end{pmatrix}.$$

Finally, (29.13) becomes

$$(29.19) \quad \psi'(x') := \begin{pmatrix} \exp(\frac{i\theta}{2}\sigma\mathbf{n}) & 0 \\ 0 & \exp(\frac{i\theta}{2}\sigma\mathbf{n}) \end{pmatrix} \psi(\hat{R}(-\theta\mathbf{n})x'), \quad x' \in \mathbb{R}^4.$$

Remarks 29.11 *i) The matrices (29.18) are unitary, because they represent pure rotations.*

ii) A map $\hat{R}(\theta\mathbf{n}) \mapsto \Gamma(\hat{R}(\theta\mathbf{n}))$ is a homomorphism of the one-parametric subgroup of the Lorentz group L , into the unitary group $U(4)$.

iii) Our calculation prove that (29.18) satisfies the identity (29.11) for $\theta \in (-\pi, \pi)$. However it holds for any $\theta \in \mathbb{R}$ by analytic continuation and gives, for $\theta = 2\pi$, that

$$(29.20) \quad \Gamma(\hat{R}(2\pi\mathbf{n})) = -1.$$

Hence, (29.19) implies that the Dirac spinor changes sign by a rotation through 2π .

Corollary 29.12 *The infinitesimal generator G of rotations about the axis \mathbf{n} satisfies the identity*

$$(29.21) \quad -i\hbar G = \frac{\hbar}{2} \begin{pmatrix} \sigma\mathbf{n} & 0 \\ 0 & \sigma\mathbf{n} \end{pmatrix} + \hat{\mathbf{L}}\mathbf{n},$$

where $\hat{\mathbf{L}}$ stands for the standard orbital angular momentum operator.

Proof The infinitesimal generator is defined as the derivative in θ of the rotation around \mathbf{n} . Formula (28.1) implies that the generator is given by

$$(29.22) \quad G\psi(x) := \left. \frac{d}{d\theta} \right|_{\theta=0} \left(\Gamma(\hat{R}(\theta\mathbf{n}))\psi(\hat{R}(-\theta\mathbf{n})x) \right), \quad x \in \mathbb{R}^4.$$

Therefore, $i\hbar G$ consists of two terms:

$$(29.23) \quad -i\hbar G\psi(x) := -i\hbar \left(\left. \frac{d}{d\theta} \right|_{\theta=0} \Gamma(\hat{R}(\theta\mathbf{n})) \right) \psi(x) - i\hbar \left. \frac{d}{d\theta} \right|_{\theta=0} \psi(\hat{R}(-\theta\mathbf{n})x).$$

Differentiating (29.18), we obtain that the first term coincides with the first summand in (29.21).

Exercise 29.13 *Check that the second term in (29.23) coincides with the second summand in (29.21).*

Hints: *i) Choose the coordinates with $e_3 = \mathbf{n}$. Then $\hat{\mathbf{L}}\mathbf{n} = \hat{\mathbf{L}}_3$ and θ is an angle of rotation around e_3 in a positive direction. ii) Use the formula*

$$(29.24) \quad \hat{\mathbf{L}}_3\psi(x) = i\hbar \frac{\partial}{\partial \theta} \psi(\hat{R}(\theta\mathbf{n})x).$$

Now (29.21) is proved. ■

Exercise 29.14 *Check the formula (29.24).*

30 Coupling to Maxwell Field

We define the Dirac equation with an external Maxwell field and prove gauge invariance.

30.1 Dirac Equation in the Maxwell Field

Let us denote $\gamma_0 = \gamma^0$, $\gamma_j = -\gamma^j$ and

$$(30.1) \quad P^0 = i\hbar \frac{\partial}{\partial x^0}, \quad P^j = -i\hbar \frac{\partial}{\partial x^j}.$$

Write the Dirac equation (24.4) in the form

$$(30.2) \quad \gamma(P)\psi(x) = \mu c\psi(x), \quad x \in \mathbb{R}^4.$$

where the differential operator $\gamma(P) := \gamma_\alpha P^\alpha$ is called the *Dirac operator*.

Let us recall the Schrödinger equation with the Maxwell field:

$$(30.3) \quad [i\hbar \partial_t - e\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [-i\hbar \nabla_{\mathbf{x}} - \frac{e}{c} \mathbf{A}(t, \mathbf{x})]^2 \psi(t, \mathbf{x}),$$

In the notations (30.1)

$$(30.4) \quad c[P^0 - \frac{e}{c}\phi(t, \mathbf{x})]\psi(t, \mathbf{x}) = \frac{1}{2\mu} [\mathbf{P} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})]^2 \psi(t, \mathbf{x}).$$

In other words, in the presence of the Maxwell field, we change P^0 to $P^0 - \frac{e}{c}\phi(t, \mathbf{x})$ and \mathbf{P} to $\mathbf{P} - \frac{e}{c}\mathbf{A}(t, \mathbf{x})$. This suggests the following generalization of the Dirac equation (30.2) for the electron field in an external Maxwell field:

$$(30.5) \quad \gamma(P - \frac{e}{c}\mathcal{A}(x))\psi(x) = \mu c\psi(x), \quad x \in \mathbb{R}^4,$$

where $\mathcal{A}(x) = (\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x}))$.

Definition 30.1 We accept (30.5) as the definition of the Dirac equation for the electron field $\psi(x)$ in the presence of an external Maxwell field $\mathcal{A}(x)$.

30.2 Gauge Invariance

Let us recall that the gauge transformation

$$(30.6) \quad \phi(t, \mathbf{x}) \mapsto \phi(t, \mathbf{x}) + \frac{1}{c} \dot{\chi}(t, \mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) \mapsto \mathbf{A}(t, \mathbf{x}) - \nabla_{\mathbf{x}} \chi(t, \mathbf{x})$$

does not change the Maxwell fields corresponding to the potentials $\phi(t, \mathbf{x})$ and $\mathbf{A}(t, \mathbf{x})$. Let us rewrite it as

$$(30.7) \quad \phi(x) \mapsto \phi'(x) := \phi(x) + \frac{1}{i\hbar} P^0 \chi(x), \quad \mathbf{A}(x) \mapsto \mathbf{A}'(x) := \mathbf{A}(x) + \frac{1}{i\hbar} \mathbf{P} \chi(x)$$

Theorem 30.2 Let $\psi(x)$ be a solution of the Dirac equation (30.5) with the potentials $\phi(x), \mathbf{A}(x)$. Then $\psi'(x) := \exp(\frac{e\chi(x)}{i\hbar c})\psi(x)$ satisfies (30.5) with the potentials $\phi'(x), \mathbf{A}'(x)$.

Proof This follows from a direct calculation, since $P \exp(\frac{e\chi(x)}{i\hbar c}) = \exp(\frac{e\chi(x)}{i\hbar c}) \frac{e}{i\hbar c} P \chi(x)$. ■

31 Pauli Equation as Nonrelativistic Approximation

We demonstrate that the Pauli Equation corresponds to the nonrelativistic approximation of the Dirac equation. This justifies the correspondence between relativistic and nonrelativistic quantum theories and the Landé factor $g = 2$, predicted by Goudsmith and Uhlenbeck, for the spin.

First, rewrite the Dirac equation (30.5), similarly to (30.4), as

$$(31.1) \quad (i\hbar\partial_t - e\phi)\psi = \gamma_0[-c\gamma(\mathbf{P} - \frac{e}{c}\mathbf{A})\psi + \mu c^2\psi],$$

where $\mathbf{g} := (\gamma_1, \gamma_2, \gamma_3) := (-\gamma^1, -\gamma^2, -\gamma^3)$. Let us substitute here the block form of the γ matrices (24.11) and the splitting of the wave function (25.3). Then the Dirac equation reduces to the coupled equations

$$(31.2) \quad \begin{cases} (i\hbar\partial_t - e\phi)\psi_+ = c\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A})\psi_- + \mu c^2\psi_+ \\ (i\hbar\partial_t - e\phi)\psi_- = c\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A})\psi_+ - \mu c^2\psi_- \end{cases}$$

Now let us replace ψ by a gauge transformation $\tilde{\psi} := \exp(i\mu c^2 t/\hbar)\psi$. This cancels the term μc^2 in the first equation and doubles it in the second one:

$$(31.3) \quad \begin{cases} (i\hbar\partial_t - e\phi)\tilde{\psi}_+ = c\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A})\tilde{\psi}_- \\ (i\hbar\partial_t - e\phi)\tilde{\psi}_- = c\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A})\tilde{\psi}_+ - 2\mu c^2\tilde{\psi}_- \end{cases}$$

Let us assume that the LHS of the last equation is small compared to μc^2 . This is true in the limit $c \rightarrow \infty$ which corresponds to the *nonrelativistic approximation*. Namely, the Lorentz transformations become then the Galilean ones, the retarded potentials becomes the Coulomb ones, etc.

Remark 31.1 *The LHS represents a kinetic energy and an electrostatic potential. Usually both are small compared to the rest energy μc^2 .*

So the last equation can be approximated by

$$(31.4) \quad 2\mu c^2\tilde{\psi}_- = c\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A})\tilde{\psi}_+.$$

Substituting it into the first equation (31.3), we obtain

$$(31.5) \quad (i\hbar\partial_t - e\phi)\tilde{\psi}_+ = \frac{1}{2\mu}(\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A}))^2\tilde{\psi}_+.$$

Now we have to evaluate the operator on the RHS.

Lemma 31.2 *The identity*

$$(31.6) \quad (\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A}))^2 = (\mathbf{P} - \frac{e}{c}\mathbf{A})^2 + \frac{e}{c}\hbar\sigma\mathbf{B},$$

holds, where $\mathbf{B} := \text{rot } \mathbf{A}$ is the magnetic field.

Proof First, we have the standard identity similar to (29.17):

$$(31.7) \quad (\sigma(\mathbf{P} - \frac{e}{c}\mathbf{A}))^2 = (\mathbf{P} - \frac{e}{c}\mathbf{A})(\mathbf{P} - \frac{e}{c}\mathbf{A}) + i\sigma((\mathbf{P} - \frac{e}{c}\mathbf{A}) \times (\mathbf{P} - \frac{e}{c}\mathbf{A})).$$

Let us note that its proof does not depend on the commutation of the components of the vector $\mathbf{P} - \frac{e}{c}\mathbf{A}$. On the other hand, the vector product of the vector $\mathbf{P} - \frac{e}{c}\mathbf{A}$ with itself does not vanish since its components do not commute. For example, let us calculate the first component:

$$(31.8) \quad (P_2 - \frac{e}{c}A_2)(P_3 - \frac{e}{c}A_3) - (P_3 - \frac{e}{c}A_3)(P_2 - \frac{e}{c}A_2) = [P_2 - \frac{e}{c}A_2, P_3 - \frac{e}{c}A_3].$$

The commutator obviously reduces to

$$(31.9) \quad -\frac{e}{c}([P_2, A_3] + [A_2, P_3]) = i\frac{e}{c\hbar}(\partial_2 A_3 - \partial_3 A_2) = i\frac{e}{c\hbar}(\text{rot } A)_1 = i\frac{e}{c\hbar}\mathbf{B}_1. \quad \blacksquare$$

Now (31.5) becomes

$$(31.10) \quad i\hbar\partial_t\tilde{\psi}_+ = \frac{1}{2\mu}\left(\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2 - \frac{e}{c}\hbar\sigma\mathbf{B}\right)\tilde{\psi}_+ + e\phi\tilde{\psi}_+$$

which coincides with the Pauli equation (19.12).

Remark 31.3 This agreement with the Pauli spin theory was one of the great triumphs of the Dirac theory. It means that the Dirac equation automatically explains the Landé factor $g = 2$ for the spinor electron magnetic moment, suggested by Goudsmith and Uhlenbeck to explain the Einstein-de Haas and Stern-Gerlach experiments.

Remark 31.4 For the Dirac equation only the sum of orbital and spinor angular momentum is conserved, while for the Pauli equation both are conserved separately. This means that the Dirac equation requires both moments for relativistic invariance, while the Pauli equation is a degenerate version of the Dirac one.

32 Charge Continuity Equation

We define the charge-current densities of the Dirac equation and prove the continuity equation.

Definition 32.1 For a spinor wave function $\psi(x)$

i) The corresponding charge-current densities are defined by

$$(32.1) \quad \left\{ \begin{array}{l} \rho(x) = e\psi^\dagger(x)(\gamma^0)^2\psi(x) := e\bar{\psi}_\nu(x)\psi_\nu(x) \\ \mathbf{j}^k(x) = e\psi^\dagger(x)\gamma^0\gamma^k\psi(x) := e\bar{\psi}_\nu(x)(\gamma^0\gamma^k\psi(x))_\nu \end{array} \right. \quad x \in \mathbb{R}^4,$$

where ψ^\dagger stands for the conjugate transpose to ψ .

ii) The vector $s(x) := (\rho(x), \mathbf{j}(x))$ with the components $s^\alpha(x) = e\psi^\dagger(x)\gamma^0\gamma^\alpha\psi(x)$ is called **four-current density**.

Remark 32.2 The charge density is non-positive and equal to

$$(32.2) \quad \rho(x) = e\psi^\dagger(x)\psi(x) := e|\psi(x)|^2.$$

Theorem 32.3 For any solution $\psi(x)$ to the Dirac equation (30.5), the four-current density satisfies the continuity equation

$$(32.3) \quad \frac{\partial s^\alpha}{\partial x^\alpha} = 0, \quad x \in \mathbb{R}^4.$$

Proof Substituting the definitions (32.1), we obtain

$$(32.4) \quad i\hbar \frac{\partial s^\alpha}{\partial x^\alpha} = i\hbar e \left[\frac{\partial}{\partial x^0} (\psi^\dagger \psi) + \frac{\partial}{\partial x^k} (\psi^\dagger \gamma^0 \gamma^k \psi) \right].$$

Each term produces two summands, with derivatives of ψ or ψ^\dagger . The contribution coming from ψ is

$$(32.5) \quad i\hbar e \psi^\dagger \gamma^0 \left[\gamma^0 \frac{\partial}{\partial x^0} + \gamma^k \frac{\partial}{\partial x^k} \right] \psi = e \psi^\dagger \gamma^0 \gamma(P) \psi = e \psi^\dagger \gamma^0 \left[\mu c + \frac{e}{c} \gamma(\mathcal{A}) \right] \psi$$

by the Dirac equation (32.1). The contribution from ψ^\dagger can be evaluated similarly, with the help of the identities $(\gamma^0 \gamma^k)^\dagger = \gamma^0 \gamma^k$: it is equal to

$$(32.6) \quad -e \left[i\hbar \gamma^0 \left(\gamma^0 \frac{\partial}{\partial x^0} + \gamma^k \frac{\partial}{\partial x^k} \right) \psi \right]^\dagger \psi = -e (\gamma^0 \gamma(P) \psi)^\dagger \psi = -e \left[\gamma^0 \left(\mu c + \frac{e}{c} \gamma(\mathcal{A}) \right) \psi \right]^\dagger \psi.$$

Finally, the sum of the contributions from ψ and ψ^\dagger vanishes by the same identities. ■

33 Charged Antiparticles

We establish a one-to-one correspondence between solutions to the Dirac equations with charge e and $-e$, respectively. Namely, consider the Dirac equation (30.5) with $-e$ instead of e :

$$(33.1) \quad \gamma(P + \frac{e}{c}\mathcal{A}(x))\psi(x) = \mu c\psi(x), \quad x \in \mathbb{R}^4,$$

It describes the positron wave field corresponding to particles with positive charge $-e = |e|$. We will establish an isomorphism between the solutions to (30.5) and (33.1).

Definition 33.1 *The charge conjugation operator K maps a Dirac wave function to $\psi_c = K\psi := \gamma^2\bar{\psi}$, and ψ_c is called the **charge-conjugated wave function**.*

Let us note that the operator K interchanges the first two components with the last two components of the Dirac spinor. Hence it changes the energy sign: the factor $\exp(-iEt)$ interchanges with $\exp(iEt)$

Theorem 33.2 *Let ψ satisfy the Dirac equation $\gamma(P - \frac{e}{c}\mathcal{A}(x))\psi(x) = \mu c\psi(x)$ for mass μ and charge e . Then ψ_c satisfies the Dirac equation (33.1) for mass μ and charge $-e$.*

Proof *Step i)* Let us check that for any vector $p \in \mathbb{C}^4$,

$$(33.2) \quad \gamma^2\gamma(\bar{p})\gamma^2 = \overline{\gamma(p)}.$$

That is, all matrices γ^α are real except γ^2 which is purely imaginary, hence

$$(33.3) \quad \overline{\gamma^\alpha} = \begin{cases} \gamma^\alpha, & \alpha \neq 2 \\ -\gamma^\alpha, & \alpha = 2. \end{cases}$$

On the other hand, the anticommutation relations for the Dirac matrices imply

$$(33.4) \quad \gamma^2\gamma^j\gamma^2 = \begin{cases} \gamma^j, & j \neq 2 \\ -\gamma^j, & j = 2. \end{cases}$$

Therefore, (33.2) follows.

Step ii) Conjugating the Dirac equation for ψ , we obtain

$$(33.5) \quad \overline{\gamma^2\gamma(P - \frac{e}{c}\mathcal{A}(x))\gamma^2\psi} = \mu c\bar{\psi}.$$

Here $\frac{e}{c}\mathcal{A}(x)$ is real, while P includes the imaginary factor i . Hence, multiplying (33.5) by γ^2 , we get

$$(33.6) \quad \gamma(P + \frac{e}{c}\mathcal{A}(x))\psi_c = \mu c\psi_c. \quad \blacksquare$$

34 Hydrogen Atom via Dirac Equation

We determine the quantum stationary states of the hydrogen atom from the Dirac equation. The main issue is the spherical symmetry of the problem. We apply the method of separation of variables and take into account the information on the irreducible representations of the rotation group from Lecture 9.

34.1 Spectral Problem and Spherical Symmetry

Let us consider the Dirac equation for the electron field in the hydrogen atom. The corresponding four-potential of the nucleus is $\mathcal{A} = (\phi, 0, 0, 0)$ with $\phi = -e/|\mathbf{x}|$. Then the corresponding Dirac equation becomes

$$(34.1) \quad i\hbar\partial_t\psi = \mathcal{H}_D\psi := c\gamma^0(\mu c - \boldsymbol{\gamma}(\mathbf{P}))\psi + e\phi(\mathbf{x})\psi.$$

We are going to determine all quantum stationary states which are the solutions of the form $\psi_E(\mathbf{x})e^{-iEt/\hbar}$ with finite charge

$$(34.2) \quad Q(\psi_E) = \int |\psi_E(\mathbf{x})|^2 d\mathbf{x} < \infty$$

(see (32.2)). Substituting into the Dirac equation (34.1), we get the corresponding stationary eigenvalue problem

$$(34.3) \quad E\psi_E = \mathcal{H}_D\psi_E.$$

It reduces to the coupled equations for the components of the spinor $\psi_E = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}$:

$$(34.4) \quad \begin{cases} (E - \mu c^2 - e\phi)\psi_+ = c\boldsymbol{\sigma}\mathbf{P}\psi_- \\ (E + \mu c^2 - e\phi)\psi_- = c\boldsymbol{\sigma}\mathbf{P}\psi_+. \end{cases}$$

The main issue for the solution of the problem is its spherical symmetry. Namely, the nucleus potentials are spherically symmetric and it is possible to prove that the angular momentum $\mathbf{J} = \mathbf{J}(\psi) = \langle \psi, \hat{\mathbf{J}}\psi \rangle$ is conserved for the solutions to (34.1). Here, by Definition 26.1,

$$(34.5) \quad \hat{\mathbf{J}} := \hat{\mathbf{L}} + \hat{\mathbf{S}}$$

is the total angular momentum operator, where $\hat{\mathbf{L}} := -i\hbar\mathbf{x} \times \nabla$ and $\hat{\mathbf{S}} := \frac{1}{2}\hbar\boldsymbol{\sigma}$ are the orbital and spin angular momentum operator, respectively. The conservation follows from the commutation

$$(34.6) \quad [\hat{\mathbf{J}}, \mathcal{H}_D] = 0.$$

Exercise 34.1 Check (34.6). **Hint:** For the free Dirac equation with $\phi = 0$, the commutation is proved in Theorem 26.2. It remains to check the commutation of $\hat{\mathbf{J}}$ with $\phi(\mathbf{x})$ which follows obviously from the spherical symmetry of the potential.

34.2 Spherical Spinors and Separation of Variables

Let us recall that we have solved the spectral problem for the nonrelativistic Schrödinger equation by a general strategy of separation of variables (Section 9.1). Now we are going to develop it analogously for the relativistic problem (34.3). In this case, the role of the orbital angular momentum \mathbf{L} is played by the total angular momentum \mathbf{J} since it is conserved. Hence, the strategy now has to be modified correspondingly:

I. First, (34.6) implies that the operator $\hat{\mathbf{J}}^2 := \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$ commutes with \mathcal{H}_D :

$$(34.7) \quad [\hat{\mathbf{J}}^2, \mathcal{H}_D] = 0.$$

Second, \mathcal{H}_D also commutes with each $\hat{\mathbf{J}}_n$. Hence, each eigenspace of the Dirac operator \mathcal{H}_D is invariant with respect to each operator $\hat{\mathbf{J}}_n$ and $\hat{\mathbf{J}}^2$. Moreover, the operator $\hat{\mathbf{J}}^2$ commutes also with each operator $\hat{\mathbf{J}}_n$, for example,

$$(34.8) \quad [\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3] = 0.$$

Exercise 34.2 Check (34.8). **Hint:** First, prove the commutation relations $[\hat{\mathbf{J}}_k, \hat{\mathbf{J}}_j] = -i\hbar\epsilon_{kjl}\hat{\mathbf{J}}_l$ where ϵ_{kjl} is a totally antisymmetric tensor. The relations follow from similar ones for the orbital and spinor angular momenta, and from the commutation of the momenta.

Hence, we could expect that there is a basis of common eigenfunctions for the operators \mathcal{H}_D , $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$. Therefore, it would be helpful to first diagonalize simultaneously $\hat{\mathbf{J}}^2$ and $\hat{\mathbf{J}}_3$.

II. The condition (34.2) means that we consider the eigenvalue problem (34.3) in the Hilbert space $\mathcal{E} := L^2(\mathbb{R}^3) \otimes \mathbb{C}^4$. On the other hand, both operators, $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$, act only on the spinor variables and angular variables in spherical coordinates. Hence, the operators act also in the Hilbert space $\mathcal{E}_1 := L^2(S, dS) \otimes \mathbb{C}^4$, where S stands for the two-dimensional sphere $|\mathbf{x}| = 1$. Moreover, both operators $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$ have the block form

$$(34.9) \quad \hat{\mathbf{J}}_3 = \begin{pmatrix} \hat{\mathbf{L}}_3 + \hat{\mathbf{s}}_3 & 0 \\ 0 & \hat{\mathbf{L}}_3 + \hat{\mathbf{s}}_3 \end{pmatrix}, \quad \hat{\mathbf{J}}^2 = \begin{pmatrix} (\hat{\mathbf{L}} + \hat{\mathbf{s}})^2 & 0 \\ 0 & (\hat{\mathbf{L}} + \hat{\mathbf{s}})^2 \end{pmatrix},$$

where $\hat{\mathbf{s}} := \frac{1}{2}\hbar\sigma$. Therefore, we can split the space \mathcal{E}_1 into the sum $\mathcal{E}_1 = E_1^+ \oplus E_1^-$, where $E_1^\pm = E_1 := L^2(S, dS) \otimes \mathbb{C}^2$, and the action of the operators is identical in each summand. In the following section we will prove

Lemma 34.3 *i) In the space E_1 there exists an orthonormal basis of Spinor Spherical Harmonics $\mathcal{S}_{jk}(\theta, \varphi)$ which are common eigenfunctions of the operators $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$:*

$$(34.10) \quad \hat{\mathbf{J}}_3 \mathcal{S}_{jk}(\theta, \varphi) = \hbar k \mathcal{S}_{jk}(\theta, \varphi), \quad \hat{\mathbf{J}}^2 \mathcal{S}_{jk}(\theta, \varphi) = \hbar^2 j(j+1) \mathcal{S}_{jk}(\theta, \varphi), \quad k = -j, -j+1, \dots, j,$$

where $j = \frac{1}{2}, \frac{3}{2}, \dots$

ii) The space of the solutions to (34.10) is two-dimensional for each fixed j, k except for a one-dimensional space for $j = 1/2, k = -1/2$.

III. The lemma suggests that we could construct the eigenfunctions of the Dirac operator \mathcal{H}_D , by separation of variables, in the form

$$(34.11) \quad \psi_E = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} R_+^+(r) \mathcal{S}_{jk}^+(\theta, \varphi) + R_+^-(r) \mathcal{S}_{jk}^-(\theta, \varphi) \\ R_-^+(r) \mathcal{S}_{jk}^+(\theta, \varphi) + R_-^-(r) \mathcal{S}_{jk}^-(\theta, \varphi) \end{pmatrix}$$

where $\mathcal{S}_{jk}^+, \mathcal{S}_{jk}^-$ is the basis of the solutions to the equations (34.10) with the fixed j, k . The following theorem justifies this particular choice for the eigenfunctions.

Lemma 34.4 *(On Separation of Variables) Each solution to the spectral problem (34.3) is a sum (or a series) of solutions of the particular form (34.11).*

Formal Proof Let π_{jk} denote the orthogonal projection in E_1 onto the linear span E_{jk} of \mathcal{S}_{jk} . Let us define its action also in $\mathcal{E} = E_1 \otimes L^2(\mathbb{R}^+) \oplus E_1 \otimes L^2(\mathbb{R}^+)$ as $\pi_{jk} \otimes 1 \oplus \pi_{jk} \otimes 1$, or equivalently,

$$(34.12) \quad [\Pi_{jk} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}](r, \cdot, \cdot) := \begin{pmatrix} \pi_{jk}[\psi_+(r, \cdot, \cdot)] \\ \pi_{jk}[\psi_-(r, \cdot, \cdot)] \end{pmatrix}, \quad r > 0.$$

in the spherical coordinates r, θ, φ . Then Π_{jk} commutes with the Schrödinger operator \mathcal{H} since the latter commutes with \mathbf{H}_3 and \mathbf{H}^2 :

$$(34.13) \quad [\mathcal{H}, \Pi_{jk}] = 0.$$

Hence, applying Π_{jk} to (34.3), we get *formally*

$$(34.14) \quad E_\omega \Pi_{jk} \psi_\omega = \mathcal{H} \Pi_{jk} \psi_\omega.$$

It remains to note that

i) The function $\Pi_{jk} \psi_\omega$ has the form (34.11) since $\pi_{jk}[\psi_\pm(r, \cdot, \cdot)] \in E_{jk}$, and $\dim \mathcal{E}_{jk} \leq 2$;

ii) $\psi_\omega = \sum_{j,k} \Pi_{jk} \psi_\omega$. ■

Remarks 34.5 *i) A complete solution to the spectral problem (34.11) relies on an investigation of all commutation relations of the operators $\hat{\mathbf{J}}_k, \hat{\mathbf{L}}_k$ and $\hat{\mathbf{S}}_k, k = 1, 2, 3$, i.e. the Lie algebra generated by them.*

ii) It still remains to determine the radial functions in (34.11). We will substitute (34.11) into the equation (34.4). This gives a radial eigenvalue problem which will be solved explicitly.

Tensor product and Clebsch-Gordan Theorem

We prove Lemma 34.3. In (34.10), we consider the action of the operator $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{s}}$ in the space $E_1 := L^2(S, dS) \otimes \mathbb{C}^2$. Let us note that the operator $\hat{\mathbf{L}}$ acts on $L^2(S, dS)$ while $\hat{\mathbf{s}}$ acts on the second factor \mathbb{C}^2 , and $\hat{\mathbf{L}}$ obviously commutes with $\hat{\mathbf{s}}$. Therefore, the operator $\hat{\mathbf{J}}$ is a generator of the tensor product of the regular and spinor representations of the rotation group $SO(3)$ since $\hat{\mathbf{L}}, \hat{\mathbf{s}}$ are the generators of the representations. Then the eigenfunctions and eigenvalues of $\hat{\mathbf{J}}^2$ can be found by the Clebsch-Gordan theorem [38].

Namely, we know the spectral decomposition of the operator $\hat{\mathbf{L}}^2$ in the space $L^2(S, dS)$:

$$(34.15) \quad L^2(S, dS) = \bigoplus_{l=0}^{\infty} L(l).$$

Here the $L(l)$ are finite-dimensional orthogonal eigenspaces of the operator $\hat{\mathbf{L}}^2$ corresponding to the eigenvalues $\hbar^2 l(l+1)$, where $l = 0, 1, \dots$. In $L(l)$ there is an orthonormal basis e_{-l}, \dots, e_l where $e_m = \hat{\mathbf{H}}_+^{m+l} e_{-l}$ (here $\hat{\mathbf{H}}_+ := \hat{\mathbf{H}}_1 + i\hat{\mathbf{H}}_2$, where $\mathbf{H}_k := \hbar^{-1} \mathbf{L}_k$) and (cf. (34.10))

$$(34.16) \quad \hat{\mathbf{L}}_3 e_{lm} = \hbar m e_{lm}, \quad \hat{\mathbf{L}}^2 e_{lm} = \hbar^2 l(l+1) m e_{lm}, \quad m = -l, \dots, l.$$

Namely, $e_{lm} = \hbar Y_l^m$, where Y_l^m are Spherical Harmonics (8.8). Similarly, in \mathbb{C}^2 there is an orthonormal basis $f_{-1/2}, f_{1/2}$ where $f_{1/2} = \hat{\mathbf{s}}_+ f_{-1/2}$ and

$$(34.17) \quad \hat{\mathbf{s}}_3 f_s = \hbar s f_s, \quad \hat{\mathbf{s}}^2 f_s = \hbar^2 s(s+1) f_s, \quad s = -1/2, 1/2.$$

Namely, $f_{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $f_{1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Therefore, we have

$$(34.18) \quad L^2(S, dS) \otimes \mathbb{C}^2 = \bigoplus_{l=0}^{\infty} L(l) \otimes \mathbb{C}^2,$$

and the tensor products $e_{lm} \otimes f_s$ with $m = -l, \dots, l$ and $s = -1/2, 1/2$ form an orthonormal basis in the space $E(l) := L(l) \otimes \mathbb{C}^2$. The relations (34.16), (34.17) imply that

$$(34.19) \quad \hat{\mathbf{J}}_3 e_{lm} \otimes f_s = \hbar(l+s) e_{lm} \otimes f_s.$$

Now let us state the Clebsch-Gordan theorem for our particular case. It is known as the "addition of angular momenta".

Lemma 34.6 For each $l = 0, 1, \dots$

i) The space $E(l)$ is an orthogonal sum of two eigenspaces $E_{\pm}(l)$ of the operator $\hat{\mathbf{J}}^2$:

$$(34.20) \quad E(l) = E_+(l) \oplus E_-(l), \quad \hat{\mathbf{J}}^2|_{E_{\pm}(l)} = \hbar^2 l_{\pm}(l_{\pm} + 1),$$

where $l_{\pm} = l \pm 1/2$ and $\dim E_{\pm}(l) = 2l_{\pm} + 1$.

ii) For $l \geq 1$, in the space $E_{\pm}(l)$ there exists a basis \mathcal{S}_k^{\pm} , $k = -l_{\pm}, -l_{\pm} + 1, \dots, l_{\pm}$, satisfying the eigenvalue problem (34.10) with $j = l_{\pm}$. For $l = 0$, the space $E_-(0) = 0$ and in $E_+(0)$ there exists a basis \mathcal{S}_k^+ , $k = -1/2, 1/2$, satisfying (34.10) with $j = 1/2$.

Exercise 34.7 Prove the lemma. **Hints:** i) Let us denote $\hat{\mathbf{J}}_+ := \hat{\mathbf{J}}_1 + i\hat{\mathbf{J}}_2$. Then $\mathcal{S}_k^{\pm} := \hat{\mathbf{J}}_+^{k+l_{\pm}} \mathcal{S}_{-l_{\pm}}^{\pm}$, $k = -l_{\pm}, -l_{\pm} + 1, \dots, l_{\pm}$.

ii) Hence, the space $E_{\pm}(l)$ is uniquely determined by $\mathcal{S}_{-l_{\pm}}^{\pm} \in E(l)$ which is an eigenvector of the operator $\hat{\mathbf{J}}_3$ with the eigenvalue $-\hbar l_{\pm}$.

iii) Obviously, $\mathcal{S}_{-l_+}^+ = e_{-l} \otimes f_{-1/2}$ and it remains to construct $\mathcal{S}_{-l_-}^-$. It is orthogonal to $E_+(l)$ and belongs to the subspace $F \subset E(l)$ which consists of all eigenvectors of the operator $\hat{\mathbf{J}}_3$, with the eigenvalue $-\hbar l_-$, in the space $E(l)$. We have to choose a nonzero vector $\mathcal{S}_{-l_-}^- \in F$ which is orthogonal to $E_+(l)$. Further, consider two cases, $l \geq 1$ and $l = 1$, separately:

$l \geq 1$. In this case the eigenspace F is the two-dimensional linear span of the vectors $e_{-l} \otimes f_{1/2}$ and $e_{-l+1} \otimes f_{-1/2}$ if $l \geq 1$. The intersection $E_+(l) \cap F$ is the one-dimensional linear span of the vector $\hat{\mathbf{J}}_+ \mathcal{S}_{-l_+}^+$. Hence, $\mathcal{S}_{-l_-}^- \in F$ is determined uniquely (up to a factor) as a vector orthogonal to $\hat{\mathbf{J}}_+ \mathcal{S}_{-l_+}^+$.

$l = 0$. Now the space F is the one-dimensional span of $e_0 \otimes f_{1/2}$ since the vector e_1 does not exist in this case. Hence, $E_-(0) = 0$.

Proof of Lemma 34.3 i) The functions $\mathcal{S}_{l_{\pm}, k}^{\pm} := \mathbf{J}_3^{k+l_{\pm}} \mathcal{S}_{-l_{\pm}}^{\pm}$, $k = -l_{\pm}, \dots, l_{\pm}$, are solutions to the problem (34.10) with $j = l_{\pm}$, and form an orthogonal basis in the space $E(l)$.

ii) The solutions are linear combinations of the orthogonal functions $\mathcal{S}_{j,k}^+$ and $\mathcal{S}_{j,k}^-$ if $(j, k) \neq (1/2, -1/2)$. Otherwise, all solutions are proportional to $\mathcal{S}_{1/2, -1/2}^+$ since the function $\mathcal{S}_{1/2, -1/2}^-$ does not exist. ■

It turns out that the spaces $E_{\pm}(l)$ are eigenspaces also for the operator $\sigma \hat{\mathbf{L}}$:

Lemma 34.8 $\sigma \hat{\mathbf{L}}$ takes the value $\hbar l$ resp. $-\hbar(l+1)$ on the space $E_+(l)$ resp. $E_-(l)$.

Proof This follows immediately from the identity

$$(34.21) \quad \sigma \hat{\mathbf{L}} = [\hat{\mathbf{L}} + \frac{1}{2} \hbar \sigma]^2 - \hat{\mathbf{L}}^2 - \frac{1}{4} \hbar^2 \sigma^2$$

since $l_{\pm}(l_{\pm} + 1) - l(l+1)$ equals either $\hbar l$ or $-\hbar(l+1)$. ■

34.3 Radial Equations

We are going to substitute the expansion (34.11) into the coupled equations (34.4) to derive ordinary differential equations for the radial functions $R_{\pm}^{\pm}(r)$. For this purpose we need an expression of the operator $\sigma\mathbf{P}$ in terms of the orbital angular momentum and related operators. The following lemma gives the necessary relations.

Lemma 34.9 *The following relations hold:*

$$(34.22) \quad \sigma\mathbf{P} = |\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{xP} + i\sigma\mathbf{L}),$$

$$(34.23) \quad \sigma\mathbf{x}(\sigma\mathbf{L} + \hbar) + (\sigma\mathbf{L} + \hbar)\sigma\mathbf{x} = 0.$$

Proof The formula for products of spin matrices gives

$$(34.24) \quad (\sigma\mathbf{x})(\sigma\mathbf{P}) = \mathbf{xP} + i\sigma(\mathbf{x} \times \mathbf{P}) = \mathbf{xP} + i\sigma\mathbf{L}.$$

Now the equation (34.22) follows on multiplying this equation on the left by $\sigma\mathbf{x}$. The equation (34.23) follows on multiplying the commutation relations $[\mathbf{L}_j, \mathbf{x}_k] = i\hbar\epsilon_{jkl}\mathbf{x}_l$ by $\sigma_j\sigma_k = 2\delta_{jk} - \sigma_k\sigma_j$ and simplifying. \blacksquare

Substituting the expression (34.22) into the equations (34.4) and using (34.23), we get

$$(34.25) \quad \begin{cases} (E - \mu c^2 - e\phi)\psi_+ &= c|\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{xP} + i\sigma\mathbf{L})\psi_- \\ &= c|\mathbf{x}|^{-2}(\mathbf{xP} - i(\sigma\mathbf{L} + \hbar))\sigma\mathbf{x}\psi_- \\ (E + \mu c^2 - e\phi)\psi_- &= c|\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{xP} + i\sigma\mathbf{L})\psi_+. \end{cases}$$

The last equation can be rewritten as

$$(34.26) \quad (E + \mu c^2 - e\phi)\sigma\mathbf{x}\psi_- = c(\mathbf{xP} + i\sigma\mathbf{L})\psi_+.$$

Together with the second equation in (34.25), this suggests the substitution $\Psi_- := \frac{\sigma\mathbf{x}}{|\mathbf{x}|}\psi_-$ and $\Psi_+ := \psi_+$. Then we get, rewriting the equations (34.25) in spherical coordinates,

$$(34.27) \quad \begin{cases} (E - \mu c^2 - e\phi)\Psi_+ &= c(-i\hbar\frac{d}{dr} - ir^{-1}(\sigma\mathbf{L} + 2\hbar))\Psi_- \\ (E + \mu c^2 - e\phi)\Psi_- &= c(-i\hbar\frac{d}{dr} + ir^{-1}\sigma\mathbf{L})\Psi_+. \end{cases}$$

By Lemma 34.4, it suffices to construct all nonzero solutions to (34.4) in the form (34.11). For example, we can assume that

$$(34.28) \quad R_+^+(r) \neq 0.$$

Let us denote by π_{jk}^+ the orthogonal projection in the space E_1 onto the linear span of the function \mathcal{S}_{jk}^+ . Denote by $\Pi_{jk}^+ := \pi_{jk}^+ \otimes 1$ the corresponding projector in the space $E_1 \otimes L^2(\mathbb{R}) = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. Then Π_{jk}^+ commutes with the operator $\sigma\hat{\mathbf{L}}$ by Lemma 34.8. Hence, applying Π_{jk}^+ to the equations (34.27), we get

$$(34.29) \quad \begin{cases} (E - \mu c^2 - e\phi)\Pi_{jk}^+\Psi_+ &= c(-i\hbar\frac{d}{dr} - ir^{-1}(\hbar l + 2\hbar))\Pi_{jk}^+\Psi_- \\ (E + \mu c^2 - e\phi)\Pi_{jk}^+\Psi_- &= c(-i\hbar\frac{d}{dr} + ir^{-1}\hbar l)\Pi_{jk}^+\Psi_+, \end{cases}$$

where $l = j - \frac{1}{2}$ by the lemma. Let us note that

$$(34.30) \quad \Pi_{jk}^+ \Psi_+ = R_1(r) \mathcal{S}_{jk}^+(\theta, \varphi), \quad \Pi_{jk}^+ \Psi_- = R_2(r) \mathcal{S}_{jk}^+(\theta, \varphi),$$

where

$$(34.31) \quad R_1(r) \equiv R_+^+(r) \neq 0$$

by (34.28). Let us denote by $R(r) := \begin{pmatrix} R_1(r) \\ R_2(r) \end{pmatrix}$ and substitute the representations (34.30) into the equations (34.29). Then (34.29) is equivalent to the following *radial equation* for the vector-function $R(r)$:

$$(34.32) \quad (E - e\phi - \mu c^2 \sigma_3) R(r) = -i\hbar \left[\left(\frac{d}{dr} + \frac{1}{r} \right) + \frac{(l+1)}{r} \sigma_3 \right] \sigma_1 R(r), \quad r > 0,$$

where σ_k are the Pauli matrices.

Corollary 34.10 *The eigenvalue problem for the Dirac equation with a spherically symmetric electrostatic potential can be written in the form*

$$(34.33) \quad (E - e\phi - \mu c^2 \sigma_3) R(r) = -i\hbar \left[\left(\frac{d}{dr} + \frac{1}{r} \right) \sigma_1 + i \frac{(l+1)}{r} \sigma_2 \right] R(r), \quad r > 0$$

which follows from (34.32) since $\sigma_3 \sigma_1 = i\sigma_2$.

34.4 Hydrogen Spectrum

Here we calculate the eigenvalues E of the problem (34.3). As in the nonrelativistic case, substitute $R(r) = e^{-\kappa r} P(r)$. Then the equation (34.32) reduces to

$$(34.34) \quad (E - e\phi - \mu c^2 \sigma_3) P(r) = -i\hbar \left[\left(\frac{d}{dr} + \frac{1}{r} - \kappa \right) \sigma_1 + i \frac{(l+1)}{r} \sigma_2 \right] P(r), \quad r > 0,$$

or equivalently, to

$$(34.35) \quad (E - \mu c^2 \sigma_3 - i\hbar \kappa \sigma_1) P(r) = -i\hbar \left[\left(\frac{d}{dr} + \frac{1}{r} \right) \sigma_1 + i \frac{(l+1)}{r} \sigma_2 + \frac{ie\phi}{\hbar} \right] P(r), \quad r > 0.$$

For the above matrix we introduce the notation

$$(34.36) \quad M \equiv \frac{i}{\hbar} (E - \mu c^2 \sigma_3 - i\hbar \kappa \sigma_1),$$

because it frequently appears in the calculations. Let us also rewrite the Coulombic potential as $e\phi = -\hbar \alpha / r$, where

$$\alpha := \frac{e^2}{\hbar} \approx \frac{1}{137}$$

is the dimensionless Sommerfeld *fine structure constant*.

We find the parameter κ from the asymptotic condition at infinity. Namely, we suggest that $R(r)$ is a “polynomial”

$$(34.37) \quad R(r) = r^\delta \sum_0^n R_k r^k$$

with an $R_n \neq 0$. Then the equation (34.35) implies that $MR_n = 0$, hence $\det M = 0$:

$$(34.38) \quad \det(E - \mu c^2 \sigma_3 - i\hbar \kappa \sigma_1) = 0.$$

This is equivalent to

$$(34.39) \quad c^2 \hbar^2 \kappa^2 = \mu^2 c^4 - E^2,$$

so, in particular, $E < \mu c^2$. We have to choose the positive root for κ to have a solution satisfying (34.2). To justify (34.37), we seek a solution in the general form

$$(34.40) \quad R(r) = r^\delta \sum_0^\infty R_k r^k$$

where we can assume that $R_0 \neq 0$ for a nontrivial solution. Substituting into (34.35), we get the equation

$$(34.41) \quad \sum_0^\infty r^{k+\delta} MR_k = \sum_0^\infty \left[(k + \delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] r^{k+\delta-1} R_k.$$

This gives the recurrence equation

$$(34.42) \quad MR_{k-1} = \left[(k + \delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] r^{k+\delta-1} R_k, \quad k = 0, 1, \dots$$

This equation with $k = 0$ implies that

$$(34.43) \quad \left[(\delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] R_0 = 0.$$

This implies the "indicial equation"

$$(34.44) \quad \det \left[(\delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] = 0$$

since $R_0 \neq 0$. It is equivalent to

$$(34.45) \quad (\delta + 1)^2 = (l + 1)^2 - \alpha^2.$$

Then $|\delta + 1| \approx l + 1$ since α is small. Therefore, we have to choose the positive root for $\delta + 1$ since for the negative root we get $-\delta \approx l + 2 \geq 2$ while $\delta > -3/2$ by the condition (34.2).

Finally, an investigation of the recurrence equation (34.42) shows that the series (34.40) should terminate by the condition (34.2) as in the case of the Schrödinger equation. Hence, we arrive at (34.37) with an $R_n \neq 0$. This implies again (34.38) and (34.39), however it is not sufficient to determine the eigenvalues E since we have the additional unknown parameter κ . Therefore, we need an additional equation which we will derive from the recurrence equation (34.42) with $k = n$:

$$(34.46) \quad MR_{n-1} = \left[(n + \delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] r^{n+\delta-1} R_n.$$

Namely, the characteristic equation for the matrix M reads,

$$(34.47) \quad (M - 2iE/c\hbar)M = 0$$

since its determinant is zero and the trace is $2iE/c\hbar$. Therefore, multiplying both sides of (34.42) by $M - 2iE/c\hbar$, we get

$$(34.48) \quad 0 = (M - 2iE/c\hbar) \left[(n + \delta + 1)\sigma_1 + i(l + 1)\sigma_2 + i\alpha \right] R_n.$$

Multiplying here the Pauli matrices, we arrive at

$$(34.49) \quad 0 = \left[2\kappa(n + \delta + 1) - 2\alpha E/c\hbar \right] R_n.$$

This gives us the new quantization condition

$$(34.50) \quad \alpha E = c\hbar(n + \delta + 1),$$

which together with the equation (34.39) determines the eigenvalues E : solving the system of equations, we get

$$(34.51) \quad E = E_{ln} = \frac{\mu c^2}{\sqrt{1 + \left[\alpha^2 / (n + \delta + 1)^2 \right]}},$$

where $\delta = \delta(l)$ is given by (34.45).

Since α is small, we can approximate the eigenvalues by the binomial expansion:

$$(34.52) \quad E_{ln} \approx \mu c^2 - \frac{\mu c^2 \alpha^2}{2(n + \delta + 1)^2}.$$

Remarks 34.11 *i)* The approximation (34.52) with $\delta = 0$ coincides with the nonrelativistic spectrum of the hydrogen atom up to the unessential additive constant μc^2 .

ii) The relativistic formula depends on the angular momentum j through $\delta = \delta(l)$, while the nonrelativistic formula does not depend on the angular momentum. This was another triumph of the Dirac theory since it corresponds to the experimental observation of the *fine structure*.

Remark 34.12 *The above analysis gives also the corresponding eigenfunctions*

$$(34.53) \quad \psi_E = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \frac{\sigma_{\mathbf{x}}}{|\mathbf{x}|} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = \frac{\sigma_{\mathbf{x}}}{|\mathbf{x}|} \begin{pmatrix} R_1(r) \mathcal{S}_{jk}^+(\theta, \varphi) \\ R_2(r) \mathcal{S}_{jk}^+(\theta, \varphi) \end{pmatrix}$$

Part VII

Mathematical Appendices

35 Newton Mechanics

We recall the Newton mechanics of one and many particles, for potential force fields. In the case of a certain symmetry of the potential we derive the corresponding conservation laws.

35.1 One Particle

Newton equation

The motion of one particle of mass $m > 0$ is governed by the Newton differential equation

$$(35.1) \quad m\ddot{\mathbf{x}}(t) = F(\mathbf{x}(t), t), \quad t \in \mathbb{R}.$$

Here $\mathbf{x}(t) \in \mathbb{R}^3$ is the particle position at time t and $F(\cdot)$ is the force field. Let us assume that $F \in C^1(\mathbb{R}^3 \times \mathbb{R}, \mathbb{R}^3)$. Then the solution $\mathbf{x}(t)$ is defined uniquely by the initial conditions $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^3$, $\dot{\mathbf{x}}(0) = \mathbf{v}_0 \in \mathbb{R}^3$ by the *main theorem of ordinary differential equations*. The solution exists for $|t| \leq \varepsilon$, where $\varepsilon > 0$ depends on the initial data $\mathbf{x}_0, \mathbf{v}_0$.

Energy conservation

Let us assume that the force field F has a *potential function* (or simply *potential*) $V(\cdot) \in C^2(\mathbb{R}^3 \times \mathbb{R})$,

$$(35.2) \quad F(\mathbf{x}, t) = -\nabla V(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^3, \quad t \in \mathbb{R}.$$

Definition 35.13 *i) $\mathcal{E} := \mathbb{R}^3 \times \mathbb{R}^3$ is the **phase space** of the Newton equation, $\mathcal{E}^+ := \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ is the **extended phase space** of the Newton equation.*

ii) The energy $E(\mathbf{x}, \mathbf{v}, t)$ is the function on the extended phase space defined by

$$(35.3) \quad E(\mathbf{x}, \mathbf{v}, t) = \frac{m\mathbf{v}^2}{2} + V(\mathbf{x}, t), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathcal{E}^+.$$

Theorem 35.14 *Let the condition (35.2) hold. Further, we assume that the potential does not depend on t ,*

$$(35.4) \quad V(\mathbf{x}, t) \equiv V(\mathbf{x}), \quad (\mathbf{x}, t) \in \mathbb{R}^3 \times \mathbb{R}.$$

Then for every solution $\mathbf{x}(t) \in C^2([t_0, t_1], \mathbb{R}^3)$ to the Newton equation, the energy is conserved,

$$(35.5) \quad E(t) := E(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \text{const}, \quad t \in [t_0, t_1].$$

Proof By the chain rule of differentiation, the Newton equation (35.1), and (35.2),

$$(35.6) \quad \dot{E}(t) = m\dot{\mathbf{x}}(t) \cdot \ddot{\mathbf{x}}(t) + \nabla V(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) = [m\ddot{\mathbf{x}}(t) + \nabla V(\mathbf{x}(t))] \cdot \dot{\mathbf{x}}(t) = 0, \quad t \in [t_0, t_1].$$

Well-posedness condition

Theorem 35.15 *Let the condition (35.4) hold, and let the potential be bounded from below by some constant $C \in \mathbb{R}$,*

$$(35.7) \quad V(\mathbf{x}) \geq C, \quad \mathbf{x} \in \mathbb{R}^3.$$

Then every solution $\mathbf{x}(t)$ to the Newton equation (35.1) exists globally in time, i.e., for all $t \in \mathbb{R}$.

Proof The energy conservation (35.5) implies that the velocity is bounded, $|\dot{\mathbf{x}}(t)| \leq \bar{v}$. Then also $|\mathbf{x}(t)|$ is bounded by $\bar{v}t + \text{const}$. This provides the existence of the global solution for all $t \in \mathbb{R}$. ■

35.2 Many Particles

Newton equations

The motion of n particles with masses $m_i > 0$ is governed by the Newton differential equation

$$(35.8) \quad m_i \ddot{\mathbf{x}}_i(t) = F_i(x(t), t), \quad t \in \mathbb{R}, \quad i = 1, \dots, n.$$

Here

- i) $\mathbf{x}_i(t) \in \mathbb{R}^3$ is the position of the i -th particle at time t , $x(t) = (\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)) \in \mathbb{R}^{3n}$ and
- ii) $F_i(x(t), t) \in \mathbb{R}^3$ is the force acting on the i -th particle.

Let us assume that the force field $F(x, t) := (F_1(x, t), \dots, F_n(x, t)) \in C^1(\mathbb{R}^{3n} \times \mathbb{R}, \mathbb{R}^{3n})$. Then the solution $x(t) = (\mathbf{x}_1(t), \dots, \mathbf{x}_n(t))$ to the system (35.8) is defined uniquely by the initial conditions $x(0) = x_0 \in \mathbb{R}^{3n}$, $\dot{x}(0) = v_0 \in \mathbb{R}^{3n}$ by the *main theorem of ordinary differential equations*. The solution exists for $|t| \leq \varepsilon$, where $\varepsilon > 0$ depends on the initial data x_0, v_0 .

Energy conservation

Let us assume that the force field $F(x, t)$ has a *potential function* (or simply *potential*) $V(\cdot) \in C^2(\mathbb{R}^{3n})$,

$$(35.9) \quad F_i(x, t) = -\nabla_{\mathbf{x}_i} V(x, t), \quad x \in \mathbb{R}^{3n}, \quad i = 1, \dots, n.$$

Definition 35.16 i) $\mathcal{E} = \mathbb{R}^{3n} \times \mathbb{R}^{3n}$ is the **phase space** of the Newton system (35.8), $\mathcal{E}^+ := \mathbb{R}^{3n} \times \mathbb{R}^{3n} \times \mathbb{R}$ is the **extended phase space** of the Newton system (35.8).

ii) The energy $E(x, v, t)$ is the function on the extended phase space defined by

$$(35.10) \quad E(x, v, t) = \sum_i \frac{m_i \mathbf{v}_i^2}{2} + V(x, t), \quad (x, v) \in \mathcal{E},$$

where $v = (\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Let us call a *trajectory* any solution $x(t)$ to the Newton system (35.8).

Theorem 35.17 Let the condition (35.9) hold. Further, we assume that the potential does not depend on t ,

$$(35.11) \quad V(x, t) \equiv V(x), \quad (x, t) \in \mathbb{R}^{3n} \times \mathbb{R}.$$

Then for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^{3n})$, the energy is conserved,

$$(35.12) \quad E(t) := E(x(t), \dot{x}(t)) = \text{const}, \quad t \in [t_0, t_1].$$

Proof By the chain rule of differentiation, the Newton system (35.8), and (35.9),

$$(35.13) \quad \begin{aligned} \dot{E}(t) &= \sum_i m_i \dot{\mathbf{x}}_i(t) \cdot \ddot{\mathbf{x}}_i(t) + \sum_i \nabla_{\mathbf{x}_i} V(x(t)) \cdot \dot{\mathbf{x}}_i(t) \\ &= \sum_i [m_i \ddot{\mathbf{x}}_i(t) + \nabla_{\mathbf{x}_i} V(x(t))] \cdot \dot{\mathbf{x}}_i(t) = 0, \quad t \in [t_0, t_1]. \end{aligned}$$

Well-posedness condition

Theorem 35.18 Let the condition (35.11) hold, and let the potential be bounded from below by some constant $C \in \mathbb{R}$:

$$(35.14) \quad V(x) \geq C, \quad x \in \mathbb{R}^{3n}.$$

Then every solution $x(t)$ to the Newton equation (35.8) exists globally in time, i.e. for all $t \in \mathbb{R}$.

Proof The energy conservation (35.12) implies that the velocity is bounded, $|\dot{x}(t)| \leq \text{const}$. Then also $|x(t)|$ is bounded by $\bar{v}t + \text{const}$. This provides the existence of the global solution for all $t \in \mathbb{R}$. ■

35.3 Symmetry Theory

The invariance of the potential V with respect to translations in time, (35.11), provides the energy conservation (35.12). Let us show that the invariance of the potential $V(x)$ with respect to some transformations of the *configuration space* $Q := \mathbb{R}^{3n}$ leads to new conservation laws.

Translation group

Let us fix a vector $\mathbf{h} \neq 0$ in \mathbb{R}^3 and consider the translations $x \mapsto x + \mathbf{h}s$ of \mathbb{R}^3 and the corresponding action in \mathbb{R}^{3n} :

$$(35.15) \quad T_s(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\mathbf{x}_1 + \mathbf{h}s, \dots, \mathbf{x}_n + \mathbf{h}s), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{3n}.$$

Definition 35.19 *The system (35.8) is invariant with respect to the translations (35.15) if*

$$(35.16) \quad V(T_s(x), t) = V(x, t), \quad (x, t) \in \mathbb{R}^{3n+1}, \quad \forall s \in \mathbb{R}.$$

Example 35.20 *The Newton system (35.8) is invariant with respect to the translations (35.15) with every $\mathbf{h} \in \mathbb{R}^3$, if the potential energy has the structure*

$$(35.17) \quad V(\mathbf{x}_1, \dots, \mathbf{x}_n, t) = W(\mathbf{x}_1 - \mathbf{x}_n, \dots, \mathbf{x}_{n-1} - \mathbf{x}_n, t), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n, t) \in \mathbb{R}^{3n+1}$$

with a function W of $3n - 2$ variables.

Definition 35.21 *i) The momentum \mathbf{p}_i of the i -th particle is the vector function on the phase space \mathcal{E} defined by*

$$(35.18) \quad \mathbf{p}_i := m_i \mathbf{v}_i \in \mathbb{R}^3, \quad (x, v) \in \mathcal{E}.$$

ii) the (total) momentum \mathbf{p} of the system (35.8) is the vector function on the phase space \mathcal{E} defined by

$$(35.19) \quad \mathbf{p} := \sum_i \mathbf{p}_i = \sum_i m_i \mathbf{v}_i \in \mathbb{R}^3, \quad (x, v) \in \mathcal{E}.$$

iii) The center of mass of the system of n particles is

$$(35.20) \quad \mathbf{X} := \frac{1}{M} \sum_i m_i \mathbf{x}_i, \quad x = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{3n},$$

where $M := \sum_i m_i$ is the total mass of the system.

Theorem 35.22 *Let (35.9) hold and the system (35.8) be invariant with respect to the translations (35.15) along a fixed vector $\mathbf{h} \in \mathbb{R}^3$. Then for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^{3n})$, the projection of the momentum $\mathbf{p}(t)$ onto \mathbf{h} is conserved,*

$$(35.21) \quad \mathbf{p}_h(t) := \mathbf{p}(t) \cdot \mathbf{h} = \text{const}, \quad t \in [t_0, t_1].$$

Proof By (35.9), (35.8), and the chain rule of differentiation,

$$(35.22) \quad \dot{\mathbf{p}}_h(t) = \sum_i m_i \ddot{\mathbf{x}}_i(t) \cdot \mathbf{h} = - \sum_i \nabla_{\mathbf{x}_i} V(x(t), t) \cdot \mathbf{h} = - \frac{d}{ds} \Big|_{s=0} V(T_s x(t), t) = 0, \quad t \in [t_0, t_1]$$

by (35.16). ■

Corollary 35.23 *Let the Newton system (35.8) be invariant with respect to the translations (35.15) along all vectors $\mathbf{h} \in \mathbb{R}^3$. Then for any trajectory, the momentum $\mathbf{p}(t)$ is conserved, $\mathbf{p}(t) = \text{const}$, and the center of mass $\mathbf{X}(t) = \sum_i m_i \mathbf{x}_i(t)/M$ moves uniformly: $\mathbf{X}(t) = \mathbf{v}t + \mathbf{X}(0)$.*

Proof Since (35.22) holds for every \mathbf{h} , we have $\dot{\mathbf{p}}(t) = 0$ and

$$\ddot{\mathbf{X}}(t) = \frac{1}{M} \sum_i m_i \ddot{\mathbf{x}}_i = 0. \quad \blacksquare$$

Rotation group

Let us fix a unit vector $\mathbf{e} \in \mathbb{R}^3$ and consider the rotation around \mathbf{e} in \mathbb{R}^3 with an angle of s radian. Let us denote by $R_{\mathbf{e}}(s) \in SO(3)$ the corresponding orthogonal matrix and define the corresponding transformation in \mathbb{R}^{3n} by

$$(35.23) \quad R_s(\mathbf{x}_1, \dots, \mathbf{x}_n) = (R_{\mathbf{e}}(s)\mathbf{x}_1, \dots, R_{\mathbf{e}}(s)\mathbf{x}_n), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{3n}.$$

Definition 35.24 *The system (35.8) is invariant with respect to the rotations (35.23) if*

$$(35.24) \quad V(R_s(x), t) = V(x, t), \quad (x, t) \in \mathbb{R}^{3n+1}, \quad \forall s \in \mathbb{R}.$$

Example 35.25 *The system (35.8) is invariant with respect to the rotations (35.15) with every $\mathbf{e} \in \mathbb{R}^3$, if the potential energy has the structure*

$$(35.25) \quad V(\mathbf{x}_1, \dots, \mathbf{x}_n, t) = W(\{|\mathbf{x}_i - \mathbf{x}_j| : 1 \leq i < j \leq n\}, t), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n, t) \in \mathbb{R}^{3n+1}.$$

Definition 35.26 *i) The angular momentum \mathbf{L}_i of the i -th particle is the vector function on the phase space \mathcal{E} defined by*

$$(35.26) \quad \mathbf{L}_i(x, v) := \mathbf{x}_i \times \mathbf{p}_i \in \mathbb{R}^3, \quad (x, v) \in \mathcal{E}.$$

ii) the angular momentum \mathbf{L} of the Newton system (35.8) is the vector function on the phase space \mathcal{E} defined by

$$(35.27) \quad \mathbf{L}(x, v) := \sum_i \mathbf{L}_i = \sum_i \mathbf{x}_i \times \mathbf{p}_i \in \mathbb{R}^3, \quad (x, v) \in \mathcal{E}.$$

Theorem 35.27 *Let the Newton system (35.8) be invariant with respect to the rotations (35.23) around a fixed vector $\mathbf{e} \in \mathbb{R}^3$. Then for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^{3n})$, the projection of the angular momentum $\mathbf{L}(x(t), \dot{x}(t))$ onto \mathbf{e} is conserved,*

$$(35.28) \quad \mathbf{L}_r(t) := \mathbf{L}(x(t), \dot{x}(t)) \cdot \mathbf{e} = \text{const}, \quad t \in [t_0, t_1].$$

Proof The differentiation gives,

$$(35.29) \quad \begin{aligned} \dot{\mathbf{L}}_{\mathbf{e}}(t) &= \left[\sum_i \dot{\mathbf{x}}_i(t) \times \mathbf{p}_i(t) + \sum_i \mathbf{x}_i(t) \times \dot{\mathbf{p}}_i(t) \right] \cdot \mathbf{e} \\ &= \left[\sum_i \dot{\mathbf{x}}_i(t) \times m_i \dot{\mathbf{x}}_i(t) + \sum_i \mathbf{x}_i(t) \times m_i \ddot{\mathbf{x}}_i(t) \right] \cdot \mathbf{e} \\ &= \sum_i [\mathbf{x}_i(t) \times m_i \ddot{\mathbf{x}}_i(t)] \cdot \mathbf{e} = \sum_i m_i \ddot{\mathbf{x}}_i(t) \cdot [\mathbf{e} \times \mathbf{x}_i(t)]. \end{aligned}$$

Therefore, the Newton system (35.8) and (35.27) imply by the chain rule of differentiation,

$$(35.30) \quad \dot{\mathbf{L}}_{\mathbf{e}}(t) = - \sum_i \nabla_{\mathbf{x}_i} V(x(t), t) \cdot [\mathbf{e} \times \mathbf{x}_i(t)] = - \frac{d}{ds} \Big|_{s=0} V(R_s x(t), t) = 0, \quad t \in [t_0, t_1].$$

by (35.24), since

$$(35.31) \quad \mathbf{e} \times \mathbf{x}_i(t) = \frac{d}{ds} \Big|_{s=0} R_{\mathbf{e}}(s)\mathbf{x}_i(t).$$

This identity follows from the fact that the vectors on both sides are orthogonal to the plane containing e and $\mathbf{x}_i(t)$, and have the same length. Indeed, the length of the LHS is $|\mathbf{x}_i(t)| \sin \alpha$, where α is the angle between e and $\mathbf{x}_i(t)$, and the length of the RHS is the radius of the circle $\{R_{\mathbf{e}}(s)\mathbf{x}_i(t) : s \in [0, 2\pi]\}$ which is equal to $|\mathbf{x}_i(t)| \sin \alpha$. ■

36 Lagrangian Mechanics

We introduce Lagrangian systems corresponding to one particle and to many particles, formulate the Hamilton least action principle, derive the Euler-Lagrange equations and check that the Newton equations are of the Euler-Lagrange form. In the case of a certain symmetry of the Lagrangian function we derive the corresponding conservation laws.

36.1 One Particle

Lagrangian function

We expose the Lagrangian form of the Newton equation (35.1) with the potential (35.2),

$$(36.1) \quad m\ddot{\mathbf{x}}(t) = -\nabla V(\mathbf{x}(t), t), \quad t \in \mathbb{R}.$$

Definition 36.1 *The Lagrangian $L(\mathbf{x}, \mathbf{v}, t)$ of the system is the following function on the extended phase space $\mathcal{E}^+ = \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ (cf. (35.3)),*

$$(36.2) \quad L(\mathbf{x}, \mathbf{v}, t) = \frac{m\mathbf{v}^2}{2} - V(\mathbf{x}, t), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathcal{E}^+.$$

Exercise 36.2 *Check that the Newton equation (36.1) can be represented in the **Euler-Lagrange** form,*

$$(36.3) \quad \frac{d}{dt}L_{\mathbf{v}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) = L_{\mathbf{x}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t), \quad t \in \mathbb{R}.$$

Let us consider more general Lagrangian systems with an arbitrary function $L(\mathbf{x}, \mathbf{v}, t)$.

Definition 36.3 *i) The Lagrangian system for one particle is the dynamical system described by the Lagrangian equation (36.3) with a function $L(\mathbf{x}, \mathbf{v}, t) \in C^2(\mathcal{E}^+)$.*

ii) The momentum of the Lagrangian system is the vector-function on the extended phase space \mathcal{E}^+ defined by

$$(36.4) \quad \mathbf{p} = L_{\mathbf{v}}(\mathbf{x}, \mathbf{v}, t), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathcal{E}^+.$$

iii) The energy of the Lagrangian system is the function on the extended phase space \mathcal{E}^+ defined by

$$(36.5) \quad E(\mathbf{x}, \mathbf{v}, t) = \mathbf{p}\mathbf{v} - L(\mathbf{x}, \mathbf{v}, t), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathcal{E}^+.$$

Example 36.4 *The Newton equation (36.1) results from the Lagrangian system with the Lagrangian functional (36.2), momentum $\mathbf{p} = m\mathbf{v}$, and energy $E = \frac{m\mathbf{v}^2}{2} + V(\mathbf{x}, t)$.*

Theorem 36.5 *Let the Lagrangian not depend on time,*

$$(36.6) \quad L(\mathbf{x}, \mathbf{v}, t) = L(\mathbf{x}, \mathbf{v}), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathcal{E}^+.$$

Then for any trajectory $\mathbf{x}(t) \in C^2([t_0, t_1], \mathbb{R}^3)$, the energy is conserved, (35.5).

Proof The differentiation of (36.5) with $\mathbf{x} = \mathbf{x}(t)$ and $\mathbf{v} = \dot{\mathbf{x}}(t)$ gives,

$$(36.7) \quad \dot{E}(t) = \dot{\mathbf{p}}\mathbf{v} + \mathbf{p}\dot{\mathbf{v}} - L_{\mathbf{x}}\dot{\mathbf{x}} - L_{\mathbf{v}}\dot{\mathbf{v}} = 0$$

by Equations (36.3) and Definition (36.4). ■

Exercise 36.6 *Calculate the momentum and the energy for the Lagrangian $L(\mathbf{x}, \mathbf{v}) = -m\sqrt{1 - \mathbf{v}^2}$.*

Action functional

Definition 36.7 $C^1 = C^1([0, \infty), \mathbb{R}^3)$ is the space of all paths in three-dimensional space.

We will consider the real-valued functionals \mathcal{F} on C^1 . By definition, \mathcal{F} is a map $C^1 \rightarrow \mathbb{R}$.

Example 36.8 $\mathcal{F}(\mathbf{x}) = \int_0^T |\dot{\mathbf{x}}(t)| dt$ is the length of the path $\mathbf{x}(\cdot) \in C^1$, $t \in [0, T]$.

Definition 36.9 The Gateau differential $D\mathcal{F}(\mathbf{x})$ is the **linear** functional $C^1 \mapsto \mathbb{R}$ defined by

$$(36.8) \quad \langle D\mathcal{F}(\mathbf{x}), \mathbf{h} \rangle = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{F}(\mathbf{x} + \varepsilon\mathbf{h}), \quad \mathbf{h}(\cdot) \in C^1$$

if the derivative on the RHS exists.

Let us fix a $T > 0$.

Definition 36.10 The action is the functional on C^1 defined by

$$(36.9) \quad S_T(\mathbf{x}(\cdot)) = \int_0^T L(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) dt, \quad \mathbf{x}(\cdot) \in C^1.$$

Note that the functional is defined on the whole of C^1 if $L(\mathbf{x}, \mathbf{v}, t) \in C(\mathcal{E}^+)$. Moreover, the functional is differentiable if $L(\mathbf{x}, \mathbf{v}, t) \in C^1(\mathcal{E}^+)$:

Lemma 36.11 The Gateau differential $DS_T(\mathbf{x})$ exists for $\mathbf{x} \in C^1$.

Proof From Definition 36.9 we get by the theorem of the differentiation of integrals,

$$(36.10) \quad \begin{aligned} \langle DS_T(\mathbf{x}), \mathbf{h} \rangle : &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_0^T L(\mathbf{x}(t) + \varepsilon\mathbf{h}(t), \dot{\mathbf{x}}(t) + \varepsilon\dot{\mathbf{h}}(t), t) dt \\ &= \int_0^T [L_{\mathbf{x}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)\mathbf{h}(t) + L_{\mathbf{v}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)\dot{\mathbf{h}}(t)] dt, \end{aligned}$$

since $L(\mathbf{x}, \mathbf{v}, t) \in C^2(\mathcal{E}^+)$ by our basic assumptions. ■

Hamilton least action principle

Let us introduce the space of *variations*.

Definition 36.12 $C^1(T) = \{\mathbf{h} \in C^1 : \mathbf{h}(0) = \mathbf{h}(T) = 0\}$.

Definition 36.13 The function $\mathbf{x} \in C^1$ satisfies the Hamilton least action principle (LAP) if, for any $T > 0$,

$$(36.11) \quad \langle DS_T(\mathbf{x}), \mathbf{h} \rangle = 0, \quad \forall \mathbf{h}(\cdot) \in C^1(T)$$

Theorem 36.14 For $\mathbf{x} \in C^2([0, \infty), \mathbb{R}^3)$ the Hamilton LAP is equivalent to the Euler-Lagrange equations (36.3) with $t \in [0, T]$.

Proof The partial integration in (36.10) gives

$$(36.12) \quad \langle DS_T(\mathbf{x}), \mathbf{h} \rangle = \int_0^T [L_{\mathbf{x}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) - \frac{d}{dt} L_{\mathbf{v}}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)] \mathbf{h}(t) dt, \quad \mathbf{h} \in C_0^1(T).$$

Therefore, (36.11) is equivalent to (36.3) by the following lemma:

Lemma 36.15 Main lemma of the calculus of variations (*du Bois-Reymond*).

Let a function $f(t) \in C[0, T]$ and $\int_0^T f(t)\mathbf{h}(t) dt = 0$ for any function $\mathbf{h}(t) \in C[0, T]$ with the boundary values $\mathbf{h}(0) = \mathbf{h}(T) = 0$. Then $f(t) = 0$, $t \in [0, T]$.

Exercise 36.16 Prove the lemma.

36.2 Many Particles

Lagrangian function

We extend the Lagrangian formalism to the Newton equations (35.8) with the potential (35.9),

$$(36.13) \quad m_i \ddot{\mathbf{x}}_i(t) = -\nabla_{\mathbf{x}_i} V(x(t), t), \quad t \in \mathbb{R}.$$

We introduce the Lagrangian $L(x, v, t)$ of the system (36.13) as the following function on the extended phase space $\mathcal{E}^+ = \mathbb{R}^{3n} \times \mathbb{R}^{3n} \times \mathbb{R}$ (cf. (36.2)),

$$(36.14) \quad L(x, v, t) = \sum_i \frac{m_i \mathbf{v}_i^2}{2} - V(x, t), \quad (x, v, t) \in \mathcal{E}^+,$$

where $v = (\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Exercise 36.17 Check that the Newton equation (36.13) can be represented in the **Euler-Lagrange** form,

$$(36.15) \quad \frac{d}{dt} L_v(x(t), \dot{x}(t), t) = L_x(x(t), \dot{x}(t), t), \quad t \in \mathbb{R}.$$

Let us consider more general Lagrangian systems with the extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$, where $N = 1, 2, \dots$, and an arbitrary function $L(x, v, t)$.

Definition 36.18 *i) The Lagrangian system in the extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$ is the dynamical system described by the equations (36.15) with a function $L(x, v, t) \in C^2(\mathcal{E})$.*

ii) The momentum of the Lagrangian system is the vector function on the extended phase space \mathcal{E}^+ defined by

$$(36.16) \quad p = L_v(x, v, t), \quad (x, v, t) \in \mathcal{E}^+.$$

iii) The energy of the Lagrangian system is the function on the phase space \mathcal{E}^+ defined by

$$(36.17) \quad E(x, v, t) = pv - L(x, v, t), \quad (x, v, t) \in \mathcal{E}^+.$$

Exercise 36.19 Check that the Newton equations (36.13) result from the Lagrangian system with the Lagrangian functional (36.14), momentum $p = (\mathbf{p}_1, \dots, \mathbf{p}_n)$ where $\mathbf{p}_i = m_i \mathbf{v}_i$, and the energy

$$(36.18) \quad E = \sum_i \frac{m_i \mathbf{v}_i^2}{2} + V(x, t).$$

Theorem 36.20 Let the Lagrangian not depend on time,

$$(36.19) \quad L(x, v, t) = L(x, v), \quad (x, v, t) \in \mathcal{E}^+.$$

Then for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^N)$, the energy is conserved, (35.12).

Proof The differentiation of (36.17) with $x = x(t)$ and $v = \dot{x}(t)$ gives,

$$(36.20) \quad \dot{E}(t) = \dot{p}v + p\dot{v} - L_x \dot{x} - L_v \dot{v} = 0$$

by Equations (36.15) and Definition (36.16). ■

Action functional

Definition 36.21 $C^1 = C^1([0, \infty), \mathbb{R}^N)$ is the space of all paths in N -dimensional space.

We will consider the real-valued functionals \mathcal{F} on C^1 . By definition, \mathcal{F} is a map $C^1 \rightarrow \mathbb{R}$.

Example 36.22 $\mathcal{F}(x) = \int_0^T |\dot{x}(t)| dt$ is the length of the path $x(\cdot) \in C^1$, $t \in [0, T]$.

Definition 36.23 The Gateau differential $D\mathcal{F}(x)$ is the **linear** functional $C^1 \mapsto \mathbb{R}$ defined by

$$(36.21) \quad \langle D\mathcal{F}(x), h \rangle = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{F}(x + \varepsilon h)$$

for $h(\cdot) \in C^1$ if the derivative on the RHS exists.

Let us fix a $T > 0$.

Definition 36.24 The action is the functional on $C^1(T)$ defined by

$$(36.22) \quad S_T(x) = \int_0^T L(x(t), \dot{x}(t)) dt, \quad x(\cdot) \in C^1.$$

Note that the functional is defined on the whole of $C^1(T)$ if $L(x, v) \in C(\mathcal{E})$. Moreover, the functional is differentiable if $L(x, v) \in C^1(\mathcal{E})$:

Lemma 36.25 The Gateau differential $DS_T(x)$ exists for $x(\cdot) \in C^1$.

Proof From Definition 36.9 we get by the theorem of the differentiation of integrals,

$$(36.23) \quad \begin{aligned} \langle DS_T(x), h \rangle &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_0^T L(x(t) + \varepsilon h(t), \dot{x}(t) + \varepsilon \dot{h}(t)) dt \\ &= \int_0^T [L_x(x(t), \dot{x}(t))h(t) + L_v(x(t), \dot{x}(t))\dot{h}(t)] dt, \end{aligned}$$

since $L(x, v) \in C^2(\mathcal{E})$ by our basic assumptions. ■

Hamilton least action principle

Let us introduce the space of *variations*.

Definition 36.26 $C^1(T) = \{h(\cdot) \in C^1 : h(0) = h(T) = 0\}$.

Definition 36.27 The function $x \in C^1$ satisfies the Hamilton least action principle (LAP) if for any $T > 0$

$$(36.24) \quad \langle DS_T(x), h \rangle = 0, \quad \forall h(\cdot) \in C^1(T).$$

Theorem 36.28 For $x \in C^2([0, \infty), \mathbb{R}^N)$ the Hamilton LAP is equivalent to the Euler-Lagrange equations (36.15).

Proof The partial integration in (36.23) gives

$$(36.25) \quad \langle DS(x), h \rangle = \int_0^T [L_x(x(t), \dot{x}(t)) - \frac{d}{dt} L_v(x(t), \dot{x}(t))] h(t) dt.$$

Therefore, (36.24) is equivalent to (36.15) by the main lemma of the calculus of variations. ■

Remark 36.29 The expositions in Sections 36.1 and 36.2 formally are almost identical.

37 Noether Theory of Invariants

We construct a Noether invariant for a system whose Lagrangian is invariant with respect to a one-parametric transformation group. This construction is applied to one particle and many particles mechanical systems.

The invariance of the Lagrangian L with respect to translation in time, (36.19), provides the conservation of energy (35.5). Let us show that the invariance of the Lagrangian with respect to some transformations of the *configuration space* $X := \mathbb{R}^N$ leads to new conservation laws.

37.1 Symmetry and Noether Theorem on Invariants

Consider a group $G = \{g\}$ of differentiable transformations $g \in C^2(X, X)$ of the configuration space X of a Lagrangian system.

Definition 37.1 G is a *symmetry group* of the Lagrangian system if the identity

$$(37.1) \quad L(g(x), dg(x)v, t) = L(x, v, t), \quad (x, v, t) \in \mathcal{E}^+, \quad \forall g \in G$$

holds, where $dg : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is the differential of g .

Let us recall the definition of the differential:

$$(37.2) \quad dg(x)v := \left. \frac{dg(X(\tau))}{d\tau} \right|_{\tau=0}$$

if $X(0) = x$ and $\dot{X}(0) = v$.

Consider a **one-parametric subgroup** $\{g_s \in G : s \in \mathbb{R}\}$ of the symmetry group G ,

$$(37.3) \quad L(g_s(x), dg_s(v), t) = L(x, v, t), \quad (x, v, t) \in \mathcal{E}^+, \quad s \in \mathbb{R}.$$

Remark 37.2 Since g_s is a one-parametric subgroup, we have $g_0 = Id$, hence

$$(37.4) \quad (g_0(x), dg_0(v)) = (x, v), \quad (x, v) \in \mathcal{E}.$$

Definition 37.3 The *Noether invariant* is a function on the extended phase space \mathcal{E}^+ , which is defined by

$$(37.5) \quad I(x, v, t) = L_v(x, v, t) \left. \frac{dg_s(x)}{ds} \right|_{s=0}, \quad (x, v, t) \in \mathcal{E}^+.$$

Theorem 37.4 (E.Noether [76]) Let $x(t) \in C^2(\mathbb{R}, \mathbb{R}^N)$ be a solution to the Euler-Lagrange equations (36.15) and $\{g_s : s \in \mathbb{R}\}$ a one-parametric symmetry group of the Lagrangian system. Then $I(t) := I(x(t), \dot{x}(t), t) = \text{const}$, $t \in \mathbb{R}$.

Proof Differentiation gives

$$(37.6) \quad \dot{I}(t) = \frac{d}{dt} L_v(x(t), \dot{x}(t), t) \left. \frac{d}{ds} \right|_{s=0} g_s(x(t)) + L_v(x(t), \dot{x}(t), t) \left. \frac{d}{dt} \frac{d}{ds} \right|_{s=0} g_s(x(t)).$$

For the first summand on the RHS, Eqns. (36.15) give

$$(37.7) \quad \frac{d}{dt} L_v(x(t), \dot{x}(t), t) \left. \frac{d}{ds} \right|_{s=0} g_s(x(t)) = L_x(x(t), \dot{x}(t), t) \left. \frac{d}{ds} \right|_{s=0} g_s(x(t)).$$

For the second summand we have

$$(37.8) \quad L_v(x(t), \dot{x}(t), t) \frac{d}{dt} \frac{d}{ds} \Big|_{s=0} g_s(x(t)) = L_v(x(t), \dot{x}(t), t) \frac{d}{ds} \Big|_{s=0} \frac{d}{dt} g_s(x(t)),$$

since $\frac{\partial^2}{\partial t \partial s} = \frac{\partial^2}{\partial s \partial t}$. At last,

$$(37.9) \quad \frac{d}{dt} g_s(x(t)) = dg_s(x(t)) \dot{x}(t)$$

by Definition (37.2) of the differential of the map g_s . Now (37.6)-(37.9) gives, by the chain rule and (37.4),

$$(37.10) \quad \begin{aligned} \dot{I}(t) &= L_x(x(t), \dot{x}(t), t) \frac{d}{ds} \Big|_{s=0} g_s(x(t)) + L_v(x(t), \dot{x}(t), t) \frac{d}{ds} \Big|_{s=0} dg_s(x(t)) \dot{x}(t) \\ &= \frac{d}{ds} \Big|_{s=0} L(g_s(x(t)), dg_s(x(t)) \dot{x}(t), t) = 0 \end{aligned}$$

according to (37.1). ■

37.2 Application to Many-Particle Systems

Let us apply the Noether theorem to a Lagrangian system of n particles, i.e., with $N = 3n$ and $x = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, where $\mathbf{x}_i \in \mathbb{R}^3$.

Translation group

Let us fix a vector $h \neq 0$ in \mathbb{R}^3 and consider the transformation T_s from (35.15). By Definition 37.1 the Lagrangian system is invariant with respect to the translations (35.15) if

$$(37.11) \quad L(T_s(x), dT_s(x)v, t) = L(x, v, t), \quad (x, v, t) \in \mathcal{E}^+, \quad \forall s \in \mathbb{R}.$$

Exercise 37.5 Check that for the translations T_s , $s \in \mathbb{R}$, the differential is given by

$$(37.12) \quad dT_s(x)v = v, \quad x, v \in \mathbb{R}^{3n}.$$

Exercise 37.6 Check that the Lagrangian system is invariant with respect to the translations (35.15) for every $h \in \mathbb{R}^3$, if the Lagrangian has the structure (cf. (35.17)),

$$(37.13) \quad L(\mathbf{x}_1, \dots, \mathbf{x}_n, v, t) = \Lambda(\mathbf{x}_1 - \mathbf{x}_n, \dots, \mathbf{x}_{n-1} - \mathbf{x}_n, v, t), \quad (\mathbf{x}_1, \dots, \mathbf{x}_n, v, t) \in \mathcal{E}^+.$$

Definition 37.7 *i)* The momentum p_i of the i -th particle of the Lagrangian system is the vector function on the space \mathcal{E}^+ defined by

$$(37.14) \quad \mathbf{p}_i(x, v, t) := L_{\mathbf{v}_i}(x, v, t) \in \mathbb{R}^3, \quad (x, v, t) \in \mathcal{E}^+.$$

ii) the momentum p of the Lagrangian system is the vector function on the space \mathcal{E}^+ defined by

$$(37.15) \quad p(x, v, t) := \sum_i \mathbf{p}_i = \sum_i L_{\mathbf{v}_i} \in \mathbb{R}^3, \quad (x, v, t) \in \mathcal{E}^+.$$

Theorem 37.8 *Let the Lagrangian system (35.8) be invariant with respect to the translations (35.15) along a fixed vector $h \in \mathbb{R}^3$. Then for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^{3n})$, the projection of the momentum $p(x(t), \dot{x}(t))$ onto h is conserved,*

$$(37.16) \quad p_h(t) := p(x(t), \dot{x}(t), t)h = \text{const}, \quad t \in [t_0, t_1].$$

Proof 1 The conservation follows from the Noether theorem for the one-parametric symmetry group $g_s = T_s$ since ph coincides with the corresponding Noether invariant. Indeed, the invariant reads

$$(37.17) \quad I := L_v \frac{d}{ds} \Big|_{s=0} T_s(x) = \sum_i L_{v_i} \frac{d}{ds} \Big|_{s=0} (\mathbf{x}_i + hs) = \sum_i \mathbf{p}_i h = ph. \quad \blacksquare$$

Proof 2 By Definition 37.7, the Euler-Lagrange equations (36.13), and the chain rule of differentiation,

$$(37.18) \quad \begin{aligned} \dot{p}_h(t) = \sum_i \dot{\mathbf{p}}_i(t)h &= - \sum_i \nabla_{\mathbf{x}_i} L(x(t), \dot{x}(t), t)h \\ &= - \frac{d}{ds} \Big|_{s=0} L(T_s(x(t)), dT_s(x(t), t)\dot{x}(t)) = 0, \quad t \in [t_0, t_1] \end{aligned}$$

by (37.12) and (37.11). \blacksquare

Rotation group

Let us fix a unit vector $e \in \mathbb{R}^3$ and consider the transformation R_s from (35.23). By Definition 37.1 the Lagrangian system is invariant with respect to the rotations (35.23) if

$$(37.19) \quad L(R_s(x), dR_s(x)v, t) = L(x, v, t), \quad (x, v, t) \in \mathcal{E}^+, \quad \forall s \in \mathbb{R}.$$

Exercise 37.9 *Check that for the rotations R_s , $s \in \mathbb{R}$, the differential is given by*

$$(37.20) \quad dR_s(x)v = R_s v, \quad v \in \mathbb{R}^{3n}.$$

Exercise 37.10 *Check that the Lagrangian system (35.8) is invariant with respect to the rotations (35.23) with every $r \in \mathbb{R}^3$, if the Lagrangian has the structure (cf. (35.25))*

$$(37.21) \quad \begin{aligned} L(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{v}_1, \dots, \mathbf{v}_n, t) &= \Lambda_1(\{|\mathbf{x}_i - \mathbf{x}_j| : 1 \leq i < j \leq n; |\mathbf{v}_i| : 1 \leq i \leq n\}, t), \\ &(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{v}_1, \dots, \mathbf{v}_n, t) \in \mathcal{E}^+. \end{aligned}$$

Definition 37.11 *i) The angular momentum \mathbf{L}_i of the i -th particle is the vector function on the space \mathcal{E}^+ defined by*

$$(37.22) \quad \mathbf{L}_i(x, v, t) := \mathbf{x}_i \times \mathbf{p}_i, \quad (x, v, t) \in \mathcal{E}^+.$$

ii) the angular momentum \mathbf{L} of the Newton system (35.8) is the vector function on the space \mathcal{E}^+ defined by

$$(37.23) \quad \mathbf{L}(x, v, t) := \sum_i \mathbf{L}_i = \sum_i \mathbf{x}_i \times \mathbf{p}_i, \quad (x, v, t) \in \mathcal{E}^+.$$

Theorem 37.12 *Let the Lagrangian system be invariant with respect to the rotations (35.23) around a fixed vector $\mathbf{e} \in \mathbb{R}^3$. Then, for any trajectory $x(t) \in C^2([t_0, t_1], \mathbb{R}^{3n})$, the projection of the angular momentum $\mathbf{L}(x(t), \dot{x}(t))$ onto \mathbf{e} is conserved,*

$$(37.24) \quad \mathbf{L}_e(t) := \mathbf{L}(x(t), \dot{x}(t))\mathbf{e} = \text{const}, \quad t \in [t_0, t_1].$$

Proof 1 The conservation follows from the Noether theorem for the one-parametric symmetry group $g_s = R_s$, since \mathbf{L}_e coincides with the corresponding Noether invariant. Indeed, the invariant reads

$$\begin{aligned}
 I &= L_v \frac{d}{ds} \Big|_{s=0} R_s(x) = \sum_i L_{v_i} \frac{d}{ds} \Big|_{s=0} (R_{\mathbf{e}}(s) \mathbf{x}_i) \\
 (37.25) \quad &= \sum_i \mathbf{p}_i(\mathbf{e} \times \mathbf{x}_i) = \sum_i \mathbf{e}(\mathbf{x}_i \times \mathbf{p}_i) = \mathbf{eL}
 \end{aligned}$$

according to (35.31). ■

Proof 2 By Definition 37.11, the Euler-Lagrange equations (36.13), and the chain rule of differentiation,

$$\begin{aligned}
 \dot{\mathbf{L}}_{\mathbf{e}}(t) &= \left[\sum_i \dot{\mathbf{x}}_i(t) \times \mathbf{p}_i(t) + \sum_i \mathbf{x}_i(t) \times \dot{\mathbf{p}}_i(t) \right] \mathbf{e} \\
 &= \left[\sum_i \dot{\mathbf{x}}_i(t) \times L_{\mathbf{v}_i}(x(t), \dot{x}(t)) + \sum_i \mathbf{x}_i(t) \times L_{\mathbf{x}_i}(x(t), \dot{x}(t)) \right] \mathbf{e} \\
 (37.26) \quad &= \sum_i (\mathbf{e} \times \dot{\mathbf{x}}_i(t)) L_{\mathbf{v}_i}(x(t), \dot{x}(t)) + \sum_i (\mathbf{e} \times \mathbf{x}_i(t)) L_{\mathbf{x}_i}(x(t), \dot{x}(t)).
 \end{aligned}$$

However, (35.31) and (37.20) imply that

$$(37.27) \quad \mathbf{e} \times \mathbf{x}_i(t) = \frac{d}{ds} \Big|_{s=0} R_{\mathbf{e}}(s) \mathbf{x}_i(t), \quad \mathbf{e} \times \dot{\mathbf{x}}_i(t) = \frac{d}{ds} \Big|_{s=0} dR_{\mathbf{e}}(s) \dot{\mathbf{x}}_i(t).$$

Therefore, (37.26) implies by the chain rule and (37.19),

$$(37.28) \quad \dot{\mathbf{L}}_{\mathbf{e}}(t) = \frac{d}{dt} L(R_s(x(t)), dR_s(x(t)) \dot{x}(t)) = 0. \quad \blacksquare$$

38 Hamilton Mechanics

38.1 Legendre Transform

We introduce the Legendre transform which translates the Euler-Lagrange equations to the Hamiltonian form and study the Hamilton-Jacobi equation. Let us consider a Lagrangian system with the extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$ and the Lagrangian functional $L \in C^2(\mathcal{E}^+)$. We will identify \mathbb{R}^N with its dual space.

Definition 38.1 *i) The Legendre transform corresponding to the Lagrangian L is the map of the extended phase space \mathcal{E}^+ into itself which is defined by $\lambda : (x, v, t) \mapsto (x, p, t)$ with $p := L_v(x, v, t)$.*

ii) The Legendre transform of the function $L(x, v, t)$ on \mathcal{E}^+ is the function on $\lambda\mathcal{E}^+$ defined by $(\Lambda L)(x, p, t) \equiv pv - L(x, v, t)$ with $(x, v, t) = \lambda^{-1}(x, p, t)$, if the Legendre map λ is a C^1 -diffeomorphism $\mathcal{E}^+ \rightarrow \lambda\mathcal{E}^+$.

Exercise 38.2 *Prove that $\lambda : \mathcal{E}^+ \rightarrow \lambda\mathcal{E}^+$ is a C^1 -diffeomorphism iff the following inequality holds: $|L_{vv}(x, v, t)| \neq 0$, $(x, v, t) \in \mathcal{E}^+$.*

Example 38.3 *The inequality holds for the Lagrangian (36.14) because then the Jacobian matrix $J := L_{vv}(x, v, t)$ is diagonal and $|J| = m_1 \dots m_n \neq 0$ since all $m_i > 0$.*

Example 38.4 $\Lambda v^2 = p^2/4$, $\Lambda v^4 = 3v^4 = 3(p/4)^{4/3}, \dots$

Exercise 38.5 *Prove that $\Lambda(\Lambda L) = L$ if $\lambda : \mathcal{E}^+ \rightarrow \lambda\mathcal{E}^+$ is a C^1 -diffeomorphism.*

Theorem 38.6 *Let the Legendre transform Λ be a C^1 -diffeomorphism $\mathcal{E}^+ \rightarrow \Lambda\mathcal{E}^+$. Then Λ transforms the Euler-Lagrange equations (36.15) into the Hamiltonian form,*

$$(38.1) \quad \dot{x}(t) = H_p(x(t), p(t), t), \quad \dot{p}(t) = -H_x(x(t), p(t), t),$$

where $H(x, p, t)$ is the Legendre transform of the Lagrangian,

$$(38.2) \quad H(x, p, t) = pv - L(x, v, t), \quad p = L_v(x, v, t).$$

Proof The first equation of (38.1) follows by differentiation of the identity $H(x, p, t) \equiv pv - L(x, v, t)$:

$$(38.3) \quad H_p = v + pv_p - L_x x_p - L_v v_p = v = \dot{x}$$

since $p = L_v$ by definition, and $x_p = 0$. The second equation of (38.1) follows from the Euler-Lagrange equation (36.15):

$$(38.4) \quad H_x = p_x v + p v_x - L_x - L_v v_x = -L_x = -\dot{p}.$$

since $p = L_v$ by definition, and $p_x = 0$. ■

Remark 38.7 *In (38.3) and (38.4) the derivatives L_x mean the derivatives with fixed v and t , but in all other terms the derivatives in x mean the derivatives with fixed p and t .*

Example 38.8 *For the Lagrangian (36.14) the energy has the form (36.18), hence $H(x, p, t) = \sum_i \frac{p_i^2}{2m_i} + V(x, t)$.*

Exercise 38.9 *Calculate the momentum, energy and the Hamilton function for the Lagrangian of the relativistic particle*

$$(38.5) \quad L(\mathbf{x}, \mathbf{v}) = -\mu c^2 \sqrt{1 - \left(\frac{\mathbf{v}}{c}\right)^2}, \quad (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3.$$

Solution:

$$(38.6) \quad \mathbf{p} := \frac{\mu \mathbf{v}}{\sqrt{1 - \left(\frac{\mathbf{v}}{c}\right)^2}}, \quad \mathbf{v} = \frac{\mathbf{p}}{\sqrt{\mu^2 + \left(\frac{\mathbf{p}}{c}\right)^2}}, \quad E = \frac{\mu c^2}{\sqrt{1 - \left(\frac{\mathbf{v}}{c}\right)^2}}, \quad H = c^2 \sqrt{\mu^2 + \left(\frac{\mathbf{p}}{c}\right)^2}.$$

38.2 Hamilton-Jacobi Equation

Let us consider the Lagrangian function $L(x, v, t) \in C^2(\mathcal{E}^+)$ with the corresponding Hamilton function $H(x, p, t) \in C^2(\mathcal{E}^+)$ on the extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$, and the Cauchy problem of the type

$$(38.7) \quad \left. \begin{aligned} -\dot{S}(t, x) &= H(x, \nabla S(t, x), t), & (t, x) \in \mathbb{R}^{N+1} \\ S|_{t=0} &= S_0(x), & x \in \mathbb{R}^N, \end{aligned} \right|$$

where $S_0(x) \in C^1(\mathbb{R}^N)$ is a given function. Let us describe the Hamilton-Jacobi method of the construction of the solution to the problem (38.7).

First, consider the corresponding Cauchy problem for the Hamilton system,

$$(38.8) \quad \left. \begin{aligned} \dot{x}(t) &= H_p(x(t), p(t), t), & \dot{p}(t) &= -H_x(x(t), p(t), t) \\ x|_{t=0} &= x_0, & p|_{t=0} &= \nabla S_0(x_0) \end{aligned} \right|$$

with $x_0 \in \mathbb{R}^N$. Let us denote the solution by $(x(t, x_0), p(t, x_0))$. The solution exists and is C^1 -smooth for small $|t|$ depending on x_0 . Let us define the function \mathcal{S} by the action integral

$$(38.9) \quad \mathcal{S}(t, x_0) = S_0(x_0) + \int_0^t L(x(s, x_0), \dot{x}(s, x_0), s) ds$$

for $x_0 \in \mathbb{R}^N$ and small $|t|$. At last, let us express x_0 in $x(t, x_0)$ for small $|t|$: this is possible since the Jacobian $x_{x_0}(t, x_0) = E$ for $t = 0$. Thus, $x_0 = x_0(t, x)$ where $x_0(t, x) \in C^1(\mathbb{R} \times \mathbb{R}^N)$, hence we can define

$$(38.10) \quad \mathcal{S}(t, x) = \mathcal{S}(t, x_0(t, x)), \quad x \in \mathbb{R}^N$$

for small $|t|$.

Theorem 38.10 *Let $T > 0$ and the map $x_0 \mapsto x(t, x_0)$ be a C^1 -diffeomorphism of \mathbb{R}^N for $t \in [0, T]$. Then the function $\mathcal{S}(t, x)$ from (38.10) is the unique solution to the Cauchy problem (38.7) for $t \in [0, T]$.*

Proof The theorem follows from the properties of the differential 1-form $\omega^1 = p dx - H dt$ in \mathbb{R}^{N+1} called the *Poincaré-Cartan integral invariant* [2].

Step i) Let us consider $x_0, x_0 + \Delta x_0 \in \mathbb{R}^N$ and $\tau, \tau + \Delta \tau \in [0, T]$. Let \mathcal{M}_τ denote the following two-dimensional submanifold in the extended phase space \mathcal{E}^+ ,

$$(38.11) \quad \mathcal{M}_\tau = \{x(t, x_0 + s\Delta x_0), x(t, x_0 + s\Delta x_0), t) : s \in [0, 1], t \in [0, \tau + s\Delta \tau]\}.$$

The boundary $\partial \mathcal{M}_\tau$ is the union $\alpha \cup \gamma_1 \cup \beta \cup \gamma_0$, where

$$(38.12) \quad \left. \begin{aligned} \alpha &:= \{(x_0 + s\Delta x_0, \nabla S_0(x_0 + s\Delta x_0), 0) : s \in [0, 1]\}, \\ \beta &:= \{(x(t + s\Delta \tau, x_0 + s\Delta x_0), p(t + s\Delta \tau, x_0 + s\Delta x_0), t + s\Delta \tau) : s \in [0, 1]\}, \\ \gamma_0 &:= \{(x(t, x_0), p(t, x_0), t) : t \in [0, \tau]\}, \\ \gamma_1 &:= \{(x(t, x_0 + \Delta x_0), p(t, x_0 + \Delta x_0), t) : t \in [0, \tau + \Delta \tau]\} \end{aligned} \right|$$

are oriented according to the increment of the parameters s, t . Therefore, by the Stokes theorem,

$$(38.13) \quad \int_{\mathcal{M}_\tau} d\omega^1 = \int_\alpha \omega^1 + \int_{\gamma_1} \omega^1 - \int_\beta \omega^1 - \int_{\gamma_0} \omega^1.$$

Step ii) The central point of the proof is the observation that the restriction of the form $d\omega^1$ onto the submanifold \mathcal{M}_τ vanishes,

$$(38.14) \quad d\omega^1|_{\mathcal{M}_\tau} = 0.$$

This follows from two facts: i) the Hamilton vector field $\mathcal{H} := (H_p, -H_x, 1)$ in the extended phase space \mathcal{E}^+ is tangent to \mathcal{M}_τ at every point, and ii) $d\omega^1(\mathcal{H}, \mathcal{V}) \equiv 0$ for every vector field \mathcal{V} in \mathcal{E}^+ . The last identity follows from the following calculations: first, $d\omega^1 = dp \wedge dx - dH \wedge dt = dp \wedge dx - (H_p dp + H_x dx) \wedge dt$ is the antisymmetric bilinear form with the matrix

$$(38.15) \quad A = \begin{pmatrix} 0 & E & H_x \\ -E & 0 & H_p \\ -H_x & -H_p & 0 \end{pmatrix},$$

where E is the $N \times N$ identity matrix. Second, $A\mathcal{H} \equiv 0$, hence

$$(38.16) \quad d\omega^1(\mathcal{H}, \mathcal{V}) = \langle A\mathcal{H}, \mathcal{V} \rangle \equiv 0.$$

Step iii) Now (38.13) reads

$$(38.17) \quad \int_\beta (pdx - Hdt) = \int_\alpha \nabla S_0(x)dx + \int_{\gamma_1} Ldt - \int_{\gamma_0} Ldt,$$

since $dt|_\alpha = 0$, $p|_\alpha = \nabla S_0(x)$ and $\omega^1|_{\gamma_i} = Ldt$. The first term on the RHS of (38.17) is equal to $S_0(x_0 + \Delta x_0) - S_0(x_0)$. Therefore, (38.17) becomes by (38.9), (38.10),

$$(38.18) \quad \begin{aligned} \int_\beta (pdx - Hdt) &= S_0(x_0 + \Delta x_0) + \int_{\gamma_1} Ldt - (S_0(x_0) + \int_{\gamma_0} Ldt) \\ &= S(x + \Delta x, \tau + \Delta\tau) - S(x, \tau), \end{aligned}$$

where $x + \Delta x = x(\tau + \Delta\tau, x_0 + \Delta x_0)$ and $x = x(\tau, x_0)$. Finally, (38.18) implies $\dot{S}(t, x) = -H(x, p, t)$ and $\nabla S(t, x) = p$, hence Equation (38.9) follows. \blacksquare

39 Theory of Noether Currents

We consider a generalization of the symmetry theory to Lagrangian fields. Let the corresponding Lagrangian density be invariant with respect to a symmetry group. Then the Noether free divergent currents and the corresponding invariants can be constructed. We prove the Noether theorem providing the construction.

39.1 Field Symmetry

Consider the one-parametric group of transformations $g_s : \mathbb{R}^4 \times \mathbb{R}^N \rightarrow \mathbb{R}^4 \times \mathbb{R}^N$ of the form

$$(39.1) \quad g_s : \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} a_s(x) \\ b_s(\psi) \end{pmatrix} \Big|_{s \in \mathbb{R}}$$

where a_s and b_s are some differentiable transformations $a_s : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ and $b_s : \mathbb{R}^N \rightarrow \mathbb{R}^N$, respectively. Let us define the corresponding transformations of the fields

$$(39.2) \quad \psi(x) \mapsto \psi_s(y) := b_s(\psi(x)).$$

This definition implies a corresponding transformation for the derivatives: by the chain rule,

$$(39.3) \quad \nabla \psi(x) \mapsto \nabla_y \psi_s(y) := \nabla_\psi b_s(\psi(x)) \nabla \psi(x) \frac{\partial x(y)}{\partial y}.$$

Remark 39.1 At $s = 0$ all transformations are identities since $g_0 = Id$ for the group g_s (cf. (37.4)):

$$(39.4) \quad (a_0 x, b_0 \psi) = (x, \psi), \quad (x, \psi) \in \mathbb{R}^4 \times \mathbb{R}^N.$$

Definition 39.2 The transformation g_s , $s \in \mathbb{R}$, is a symmetry of a Lagrangian field with the Lagrangian density $\mathcal{L}(x, \psi, \nabla \psi)$ if the following identity holds,

$$(39.5) \quad \mathcal{L}(x, \psi, \nabla \psi) = \mathcal{L}(y, \psi_s, \nabla_y \psi_s) \Big|_{\frac{\partial y(x)}{\partial x}}, \quad (x, \psi, \nabla \psi) \in \mathbb{R}^{d+1} \times \mathbb{R}^N \times \mathbb{R}^{4N},$$

where $y := a_s(x)$, $\psi_s := b_s(\psi)$ and $\nabla_y \psi_s := \nabla_\psi b_s(\psi) \nabla \psi \frac{\partial x(y)}{\partial y}$.

Example 39.3 Time-translations Consider the time translation along $e_0 = (1, 0, 0, 0)$,

$$(39.6) \quad g_s : \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} x - se_0 \\ \psi \end{pmatrix} \Big|_{s \in \mathbb{R}}.$$

Then

$$(39.7) \quad \psi_s(y) = \psi(y + se_0), \quad \nabla_y \psi_s(y) = (\nabla \psi)(y + se_0), \quad \Big|_{\frac{\partial y(x)}{\partial x}} \equiv 1.$$

Hence, (39.5) for this transformation is equivalent to (3.5).

Example 39.4 Space-translations Consider the space-translation along $e_1 = (0, 1, 0, 0)$,

$$(39.8) \quad g_s : \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} x + se_1 \\ \psi \end{pmatrix} \Big|_{s \in \mathbb{R}}.$$

Then

$$(39.9) \quad \psi_s(y) = \psi(y - se_1), \quad \nabla_y \psi_s(y) = (\nabla \psi)(y - se_1), \quad \Big|_{\frac{\partial y(x)}{\partial x}} \equiv 1.$$

Hence, (39.5) for this transformation means that the Lagrangian density \mathcal{L} does not depend on \mathbf{x}_1 .

Example 39.5 Space-rotations Consider the group

$$(39.10) \quad g_s : \begin{pmatrix} (x_0, \mathbf{x}) \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} (x_0, R_n(-s)\mathbf{x}) \\ \psi \end{pmatrix} \Big| \quad s \in \mathbb{R},$$

where $R_n(s)$ is the rotation of \mathbb{R}^3 around \mathbf{e}_n with an angle of s radian. Then

$$(39.11) \quad \psi_s(y) = \psi(y_0, R_n(s)\mathbf{y}), \quad \nabla_y \psi_s(y) = (\nabla \psi) \begin{pmatrix} 1 & 0 \\ 0 & R_n(s) \end{pmatrix}, \quad \left| \frac{\partial y(x)}{\partial x} \right| \equiv 1.$$

Hence, (39.5) for this transformation is equivalent to (3.21).

Example 39.6 Phase rotations For $\psi \in \mathbb{C}^M$ define

$$(39.12) \quad g_s : \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} x \\ e^{is}\psi \end{pmatrix} \Big| \quad s \in \mathbb{R}.$$

Then

$$(39.13) \quad \psi_s(y) = e^{is}\psi(y), \quad \nabla_y \psi_s(y) = e^{is}\nabla \psi(y), \quad \left| \frac{\partial y(x)}{\partial x} \right| \equiv 1.$$

Hence, (39.5) for this transformation is equivalent to (3.24).

39.2 Noether Current and Continuity Equation

Definition 39.7 For a given one-parametric group g_s of transformations (39.1) and a given trajectory $\psi(x)$, let us define the vector fields

$$(39.14) \quad v(x) = \left. \frac{\partial a_s x}{\partial s} \right|_{s=0}, \quad w(x) = \left. \frac{\partial \psi_s(x)}{\partial s} \right|_{s=0}, \quad x \in \mathbb{R}^{d+1}.$$

Definition 39.8 The Noether current corresponding to a given one-parametric group g_s of transformations (39.1) is the following vector field

$$(39.15) \quad S_\alpha(x) = \pi_\alpha(x)w(x) + \mathcal{L}(x, \psi(x), \nabla \psi(x))v_\alpha(x), \quad x \in \mathbb{R}^{d+1}, \quad \alpha = 0, \dots, 3.$$

Theorem 39.9 (E.Noether [76]) Let g_s be a one-parametric symmetry group, i.e. (39.5) holds for $s \in \mathbb{R}$. Let $\psi(x) \in C^2(\mathbb{R}^4, \mathbb{R}^N)$ be a solution to the equations (2.9), and $w(x) \in C^1(\mathbb{R}^4, \mathbb{R}^N)$, $v(x) \in C^1(\mathbb{R}^4, \mathbb{R}^4)$ are defined by (39.14). Then the continuity equation holds,

$$(39.16) \quad \partial_\alpha S_\alpha(x) = 0, \quad x \in \mathbb{R}^4.$$

Corollary 39.10 Let i) all conditions of Theorem 39.9 hold, ii) the bounds (2.20) hold, iii) $\psi(x) \in C_\sigma^2$, $w(x) \in C_\sigma^1$ with a $\sigma > 3/2$, and $v(x) \in C_0^1$. Then the conservation law holds,

$$(39.17) \quad \mathcal{S}_0(t) := \int_{\mathbb{R}^3} S_0(t, \mathbf{x}) d\mathbf{x} = \text{const}, \quad t := x_0 \in \mathbb{R}.$$

Proof We have $\mathcal{S}_0(t) = \lim_{R \rightarrow \infty} \mathcal{S}_0^R(t)$ where

$$(39.18) \quad \mathcal{S}_0^R(t) := \int_{|\mathbf{x}| \leq R} S_0(t, \mathbf{x}) d\mathbf{x}, \quad t \in \mathbb{R}.$$

Differentiating, we get by (39.16) and the Stokes theorem,

$$(39.19) \quad \dot{\mathcal{S}}_0^R(t) := - \int_{|\mathbf{x}| \leq R} \nabla_\alpha S_\alpha(t, \mathbf{x}) d\mathbf{x} = - \int_{|\mathbf{x}|=R} n_\alpha(\mathbf{x}) S_\alpha(t, \mathbf{x}) d\Sigma, \quad t \in \mathbb{R},$$

where $n_\alpha(\mathbf{x}) := x_\alpha/|\mathbf{x}|$ and $d\Sigma$ is the Lebesgue measure on the sphere $|\mathbf{x}| = R$. Therefore, $\dot{\mathcal{S}}_0^R(t) \rightarrow 0$ as $R \rightarrow \infty$ since $S_\alpha(x) \in C_\sigma^1$ with $\sigma > 3$. Hence (39.17) follows. \blacksquare

Remark 39.11 The integral identity (39.19) means that the vector field $S_k(t, \mathbf{x})$ is the current density of the field $S_0(t, \mathbf{x})$.

Proof of Theorem 39.9 (cf. [37, 103]) Consider an arbitrary open region $\Omega \subset \mathbb{R}^3$ with a smooth boundary. Integrating the symmetry condition (39.5) over Ω , we get

$$(39.20) \quad \int_{\Omega} \mathcal{L}(x, \psi(x), \nabla \psi(x)) dx = \int_{\Omega_s} \mathcal{L}(y, \psi_s(y), \nabla_y \psi_s(y)) dy, \quad s \in \mathbb{R},$$

where $\Omega_s := a_s(\Omega)$. Let us make the change of variables $y = a_s(x)$ on the RHS. Then we get the identity

$$(39.21) \quad \int_{\Omega} \mathcal{L}(a_s(x), b_s(\psi(x)), D_s(x)) I_s(x) dx = \text{const}, \quad s \in \mathbb{R},$$

where $I_s(x) := \left| \frac{\partial a_s(x)}{\partial x} \right|$ and

$$(39.22) \quad [D_s(x)]_{\alpha} := \left. \frac{\partial \psi_s(y)}{\partial y_{\alpha}} \right|_{y=a_s(x)}, \quad \alpha = 0, \dots, d.$$

Differentiating (39.21) in s , we get by (39.4),

$$(39.23) \quad \int_{\Omega} \left[\mathcal{L}_x(x, \psi(x), \nabla \psi(x)) \cdot \left. \frac{d}{ds} \right|_{s=0} a_s(x) + \mathcal{L}_{\psi}(x, \psi(x), \nabla \psi(x)) \cdot \left. \frac{d}{ds} \right|_{s=0} b_s(\psi(x)) \right. \\ \left. + \mathcal{L}_{\nabla_{\alpha} \psi}(x, \psi(x), \nabla \psi(x)) \cdot \left. \frac{d}{ds} \right|_{s=0} [D_s(x)]_{\alpha} + \mathcal{L}(x, \psi(x), \nabla \psi(x)) \left. \frac{d}{ds} \right|_{s=0} I_s(\psi(x)) \right] dx = 0.$$

Let us calculate the four derivatives in s .

i) By Definition (39.14),

$$(39.24) \quad \left. \frac{d}{ds} \right|_{s=0} a_s(x) = v(x).$$

ii) By Definition (39.2), the chain rule and (39.14), (39.4),

$$(39.25) \quad \left. \frac{d}{ds} \right|_{s=0} b_s(\psi(x)) = \left. \frac{d}{ds} \right|_{s=0} \psi_s(a_s x) = \left. \frac{d}{ds} \right|_{s=0} [\psi_s(a_0 x) + \psi_0(a_s x)] = w(x) + \nabla \psi(x) \cdot v(x).$$

iii) By definition, $[D_s(x)]_{\alpha} := \frac{\partial b_s(\psi(x))}{\partial [a_s(x)]_{\alpha}} = \frac{\partial \psi_s(a_s(x))}{\partial [a_s(x)]_{\alpha}}$. Hence the same arguments imply

$$(39.26) \quad \left. \frac{d}{ds} \right|_{s=0} [D_s(x)]_{\alpha} = \left. \frac{d}{ds} \right|_{s=0} \frac{\partial \psi_s(a_s(x))}{\partial [a_s(x)]_{\alpha}} \\ = \left. \frac{d}{ds} \right|_{s=0} \left[\frac{\partial \psi_s(a_0(x))}{\partial [a_0(x)]_{\alpha}} + \frac{\partial \psi_0(a_s(x))}{\partial [a_0(x)]_{\alpha}} + \frac{\partial \psi_0(a_0(x))}{\partial [a_s(x)]_{\alpha}} \right] \\ = \nabla_{\alpha} w(x) + \nabla_{\alpha} (\nabla \psi(x) v(x)) + \left. \frac{d}{ds} \right|_{s=0} \frac{\partial \psi(x)}{\partial [a_s(x)]_{\alpha}}.$$

To calculate the last derivative, let us use the matrix identity

$$(39.27) \quad \frac{\partial \psi(x)}{\partial [a_s(x)]_{\beta}} \frac{\partial [a_s(x)]_{\beta}}{\partial x_{\alpha}} = \frac{\partial \psi(x)}{\partial x_{\alpha}}.$$

Differentiating in s , we get by (39.4),

$$(39.28) \quad \frac{d}{ds} \Big|_{s=0} \frac{\partial \psi(x)}{\partial [a_s(x)]_\alpha} + \frac{\partial \psi(x)}{\partial x_\beta} \frac{\partial v_\beta}{\partial x_\alpha} = 0.$$

Therefore, the last derivative in (39.26) equals $-\nabla \psi \nabla_\alpha v$. Hence (39.26) becomes,

$$(39.29) \quad \frac{d}{ds} \Big|_{s=0} [D_s(x)]_\alpha = \nabla_\alpha w(x) + \nabla(\nabla_\alpha \psi(x)) \cdot v(x).$$

iv) Finally, the derivative of the determinant I_s of the Jacobian matrix is the trace of the derivative,

$$(39.30) \quad \frac{d}{ds} \Big|_{s=0} I_s(\psi(x)) = \frac{d}{ds} \Big|_{s=0} \left| \frac{\partial a_s(x)}{\partial x} \right| = \text{tr} \frac{d}{ds} \Big|_{s=0} \frac{\partial a_s(x)}{\partial x} = \text{tr} \frac{\partial v(x)}{\partial x} = \nabla \cdot v(x),$$

since the Jacobian matrix is diagonal: $\frac{\partial a_0(x)}{\partial x} = E$.

Collecting all calculations i) – iv) in (39.23), we get

$$(39.31) \quad \int_{\Omega} \left[\mathcal{L}_x(x, \psi(x), \nabla \psi(x)) \cdot v(x) + \mathcal{L}_\psi(x, \psi(x), \nabla \psi(x)) \cdot (w(x) + \nabla \psi(x) \cdot v(x)) \right. \\ \left. + \pi_\alpha(x) \cdot (\nabla_\alpha w(x) + \nabla(\nabla_\alpha \psi(x)) \cdot v(x)) \right. \\ \left. + \mathcal{L}(x, \psi(x), \nabla \psi(x)) \nabla \cdot v(x) \right] dx = 0.$$

Since the region Ω is arbitrary, the integrand is zero by the main lemma of the calculus of variations. We rewrite it as follows,

$$(39.32) \quad \mathcal{L}_\psi \cdot w(x) + \pi_\alpha(x) \cdot \nabla_\alpha w(x) + \nabla \cdot \left[\mathcal{L}(x, \psi(x), \nabla \psi(x)) v(x) \right] = 0.$$

Finally, let us substitute $\mathcal{L}_\psi = \nabla_\alpha \pi_\alpha(x)$ from the Euler-Lagrange equations (2.9). Then (39.32) becomes

$$(39.33) \quad \nabla_\alpha \left[\pi_\alpha(x) \cdot w(x) \right] + \nabla \cdot \left[\mathcal{L}(x, \psi(x), \nabla \psi(x)) v(x) \right] = 0,$$

which coincides with (39.16) by (39.15). ■

Remark 39.12 *The justification of the formal proof (39.26) – (39.28) of (39.29) follows from the identity of type (39.27),*

$$(39.34) \quad [D_s(x)]_\beta \frac{\partial [a_s(x)]_\beta}{\partial x_\alpha} = \frac{\partial \psi(a_s x)}{\partial x_\alpha},$$

by differentiation similar to (39.28) and (39.26).

40 Application to Four Conservation Laws

We obtain the formulas for Noether currents corresponding to concrete symmetry groups: time- and space translations, space- and phase rotations. The formulas imply, by the Noether theorem, the corresponding conservation laws for energy, momentum, angular momentum and charge. We consider first general Lagrangian fields and then specify the formulas for the Klein-Gordon and Schrödinger equations.

40.1 General Lagrangian Fields

Let us apply the Noether theorem 39.9 to the four groups of the examples 39.3 – 39.6.

I. Proof of Theorem 3.4 For the group (39.6), Definition (39.14) implies by (39.7),

$$(40.1) \quad v(x) = -e_0, \quad w(x) = \nabla\psi(x)e_0 = \nabla_0\psi(x), \quad x \in \mathbb{R}^4.$$

Hence, the Noether current (39.15) becomes,

$$(40.2) \quad \begin{cases} S_0(x) &= \pi_0(x)\nabla_0\psi(x) - \mathcal{L}(x, \psi(x), \nabla\psi(x)), \\ S_k(x) &= \pi_k(x)\nabla_0\psi(x), \quad k = 1, 2, 3. \end{cases}$$

The identity (3.5) implies that the Lagrangian density satisfies the invariance condition (39.5) with the group (39.6). Therefore, Theorem 39.9 implies the continuity equation (39.16) for the current (40.2), and Corollary 39.10 implies (39.17), which means the conservation of energy by Definition (3.1).

II. Proof of Theorem 3.8 Let us consider the case $n = 1$ for concreteness. For the group (39.8), Definition (39.14) implies by (39.9),

$$(40.3) \quad v(x) = e_1, \quad w(x) = -\nabla\psi(x)e_1 = -\nabla_1\psi(x), \quad x \in \mathbb{R}^4.$$

Hence, the Noether current (39.15) becomes,

$$(40.4) \quad \begin{cases} S_0(x) &= -\pi_0(x)\nabla_1\psi(x), \\ S_1(x) &= -\pi_1(x)\nabla_1\psi(x) + \mathcal{L}(x, \psi(x), \nabla\psi(x)), \\ S_k(x) &= -\pi_k(x)\nabla_1\psi(x), \quad k = 2, 3. \end{cases}$$

The Lagrangian density satisfies the invariance condition (39.5) with the group (39.8). Therefore, Theorem 39.9 implies the continuity equation (39.16) for the current (40.4), and Corollary 39.10 implies (39.17), which means the conservation of the first component of momentum by Definition (3.16).

III. Proof of Theorem 3.13 For the group (39.10), Definition (39.14) implies by (39.11) and (35.31),

$$(40.5) \quad v(x) = (0, \mathbf{e}_n \times \mathbf{x}), \quad w(x) = \nabla\psi(x)(0, \mathbf{e}_n \times \mathbf{x}) = (\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x), \quad x \in \mathbb{R}^4.$$

Hence, the Noether current (39.15) becomes,

$$(40.6) \quad \begin{cases} S_0(x) &= \pi_0(x)(\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x), \\ S_k(x) &= \pi_k(x)(\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x) + \mathcal{L}\mathbf{e}_n \times \mathbf{x}, \quad k = 1, 2, 3. \end{cases}$$

The identity (3.21) implies that the Lagrangian density satisfies the invariance condition (39.5) with the group (39.10). Therefore, Theorem 39.9 implies the continuity equation (39.16) for the current (40.6), and Corollary 39.10 implies (39.17), which means the conservation of the n -th component of

angular momentum by Definition (3.23).

IV. Proof of Theorem 3.18 For the group (39.12), Definition (39.14) implies by (39.13),

$$(40.7) \quad v(x) = 0, \quad w(x) = i\psi(x), \quad x \in \mathbb{R}^4.$$

Hence, the components of the Noether current (39.15) equal the charge-current densities,

$$(40.8) \quad S_\alpha(x) = \pi_\alpha(x) \cdot i\psi(x), \quad \alpha = 0, \dots, 3.$$

The identity (3.24) implies that the Lagrangian density satisfies the invariance condition (39.5) with the group (39.12). Therefore, Theorem 39.9 implies the continuity equation (39.16) for the current (40.8), and Corollary 39.10 implies (39.17), which means the conservation of charge by Definition (3.25).

40.2 Klein-Gordon Equation

Let us substitute Expressions (2.16) and (2.5) into (40.2), (40.4), (40.6) and (40.8). Then Theorem 39.9 implies for solutions to Equation (2.1):

I. Energy flux The continuity equation (39.16) holds for the energy- and energy current densities

$$(40.9) \quad \left\{ \begin{array}{l} S_0(x) = -i\hbar(i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x) \cdot \nabla_0\psi(x) - \mathcal{L}(x, \psi(x), \nabla\psi(x)) \\ \\ = \frac{(i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x) \cdot (i\hbar\nabla_0 + \frac{e}{c}\phi(x))\psi(x)}{2}, \\ \\ + \sum_{k=1}^3 \frac{|(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x)|^2}{2}, \\ \\ S_k(x) = -i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot \nabla_0\psi(x), \quad k = 1, 2, 3, \end{array} \right.$$

if the potentials $\phi(x)$, $\mathbf{A}(x)$ do not depend on time $x_0 = ct$.

For the free equation (2.10):

$$(40.10) \quad \left\{ \begin{array}{l} S_0(x) = \frac{|\nabla_0\psi(x)|^2}{2} + \sum_{k=1}^3 \frac{|\nabla_k\psi(x)|^2}{2} + \mu^2 \frac{|\psi(x)|^2}{2}, \\ \\ S_k(x) = -\nabla_k\psi(x) \cdot \nabla_0\psi(x), \quad k = 1, 2, 3. \end{array} \right.$$

II. Momentum flux The continuity equation (39.16) holds for the first components of the momentum- and momentum current densities

$$(40.11) \quad \left\{ \begin{array}{l} S_0(x) = i\hbar(i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x) \cdot \nabla_1\psi(x), \\ S_1(x) = i\hbar(-i\hbar\nabla_1 - \frac{e}{c}\mathbf{A}_1(x))\psi(x) \cdot \nabla_1\psi(x) + \mathcal{L}(x, \psi(x), \nabla\psi(x)) \\ = \frac{|(-i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x)|^2}{2} + \frac{(-i\hbar\nabla_1 - \frac{e}{c}\mathbf{A}_1(x))\psi(x) \cdot (-i\hbar\nabla_1 + \frac{e}{c}\mathbf{A}_1(x))\psi(x)}{2} \\ + \sum_{k=2}^3 \frac{|(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x)|^2}{2}, \\ S_k(x) = i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot \nabla_1\psi(x), \quad k = 2, 3, \end{array} \right.$$

if the potentials $\phi(x)$, $\mathbf{A}(x)$ do not depend on x_1 .

For the free equation (2.10):

$$(40.12) \quad \left\{ \begin{array}{l} S_0(x) = -\nabla_0\psi(x) \cdot \nabla_1\psi(x), \\ S_1(x) = \frac{|\nabla_0\psi(x)|^2}{2} + \frac{|\nabla_1\psi(x)|^2}{2} - \sum_{k=2,3} \frac{|\nabla_k\psi(x)|^2}{2} - \mu^2 \frac{|\psi(x)|^2}{2}, \\ S_k(x) = \nabla_k\psi(x) \cdot \nabla_1\psi(x), \quad k = 2, 3. \end{array} \right.$$

III. Space rotations The continuity equation (39.16) holds for the n -th component of the angular momentum- and angular momentum current densities

$$(40.13) \quad \left\{ \begin{array}{l} S_0(x) = -i\hbar(i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x), \\ S_k(x) = -i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x) + \mathcal{L}(\mathbf{e}_n \times \mathbf{x})_k, \quad k = 1, 2, 3, \end{array} \right.$$

if Eq. (3.22) holds for the density (2.5).

For the free equation (2.10):

$$(40.14) \quad \left\{ \begin{array}{l} S_0(x) = \nabla_0\psi(x)(\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x), \\ S_k(x) = -\nabla_k\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n\psi(x) + \mathcal{L}(\mathbf{e}_n \times \mathbf{x})_k, \quad k = 1, 2, 3. \end{array} \right.$$

IV. Phase rotations The continuity equation (39.16) holds for the charge- and charge current densities

$$(40.15) \quad \left\{ \begin{array}{l} S_0(x) = -i\hbar(i\hbar\nabla_0 - \frac{e}{c}\phi(x))\psi(x) \cdot i\psi(x), \\ S_k(x) = -i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot i\psi(x), \quad k = 1, 2, 3. \end{array} \right.$$

For the free equation (2.10):

$$(40.16) \quad S_0(x) = \nabla_0\psi(x) \cdot i\psi(x), \quad S_k(x) = -\nabla_k\psi(x) \cdot i\psi(x), \quad k = 1, 2, 3.$$

40.3 Schrödinger Equation

Let us substitute Expressions (2.17) and (2.6) into (40.2), (40.4), (40.6) and (40.8). Then Theorem 39.9 implies for solutions to Equation (2.2):

I. Energy flux The continuity equation (39.16) holds for the energy- and energy current densities

$$(40.17) \quad \left\{ \begin{array}{l} S_0(x) = -i\hbar\psi(x) \cdot \nabla_0\psi(x) - \mathcal{L}(x, \psi(x), \nabla\psi(x)) \\ \\ = e\phi(x)\psi(x) \cdot \psi(x) + \frac{1}{2\mu} \sum_{k=1}^3 |(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x)|^2 \\ \\ S_k(x) = -\frac{1}{m}i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot \nabla_0\psi(x), \quad k = 1, 2, 3, \end{array} \right.$$

if the potentials $\phi(x)$, $\mathbf{A}(x)$ do not depend on time $x_0 = t$.

For the free equation (2.11):

$$(40.18) \quad \left\{ \begin{array}{l} S_0(x) = \frac{1}{2\mu} \sum_{k=1}^3 |\nabla_k\psi(x)|^2, \\ \\ S_k(x) = -\frac{1}{m} \nabla_k\psi(x) \cdot \nabla_0\psi(x), \quad k = 1, 2, 3. \end{array} \right.$$

II. Momentum flux The continuity equation (39.16) holds for the first component of the momentum- and momentum current densities

$$(40.19) \quad \left\{ \begin{array}{l} S_0(x) = i\hbar\psi(x) \cdot \nabla_1\psi(x), \\ \\ S_1(x) = \frac{1}{m}i\hbar(-i\hbar\nabla_1 - \frac{e}{c}\mathbf{A}_1(x))\psi(x) \cdot \nabla_1\psi(x) + \mathcal{L}(x, \psi(x), \nabla\psi(x)) \\ \\ = e\phi(x)\psi(x) \cdot \psi(x) + \frac{1}{2\mu}(-i\hbar\nabla_1 - \frac{e}{c}\mathbf{A}_1(x))\psi(x) \cdot (-i\hbar\nabla_1 + \frac{e}{c}\mathbf{A}_1(x))\psi(x) \\ \\ - \frac{1}{2\mu} \sum_{k=2}^3 |(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x)|^2, \\ \\ S_k(x) = \frac{1}{m}i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot \nabla_1\psi(x), \quad k = 2, 3, \end{array} \right.$$

if the potentials $\phi(x)$, $\mathbf{A}(x)$ do not depend on x_1 .

For the free equation (2.11):

$$(40.20) \quad \left\{ \begin{array}{l} S_0(x) = i\psi(x) \cdot \nabla_1(x), \\ \\ S_1(x) = -i\nabla_0(x) \cdot \psi(x) + \frac{1}{2\mu} |\nabla_1\psi(x)|^2 - \frac{1}{2\mu} \sum_{k=2,3} |\nabla_k\psi(x)|^2, \\ \\ S_k(x) = \frac{1}{m} \nabla_k\psi(x) \cdot \nabla_1\psi(x), \quad k = 2, 3. \end{array} \right.$$

III. Space rotations The continuity equation (39.16) holds for the n -th component of the angular momentum- and angular momentum current densities

$$(40.21) \quad \begin{cases} S_0(x) = -i\hbar\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x), \\ S_k(x) = -\frac{1}{m}i\hbar(-i\hbar\nabla_k - \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x) + \mathcal{L}(\mathbf{e}_n \times \mathbf{x})_k, \quad k = 1, 2, 3 \end{cases}$$

if Eq. (3.22) holds for the density (2.6).

For the free equation (2.11):

$$(40.22) \quad \begin{cases} S_0(x) = -i\psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x), \\ S_k(x) = -\frac{1}{m}\nabla_k \psi(x) \cdot (\mathbf{x} \times \nabla_{\mathbf{x}})_n \psi(x) + \mathcal{L}(\mathbf{e}_n \times \mathbf{x})_k, \quad k = 1, 2, 3. \end{cases}$$

IV. Phase rotations The continuity equation (39.16) holds for the charge- and charge current densities

$$(40.23) \quad S_0(x) = -\hbar\psi(x) \cdot \psi(x), \quad S_k(x) = \frac{1}{m}\hbar(i\hbar\nabla_k + \frac{e}{c}\mathbf{A}_k(x))\psi(x) \cdot \psi(x), \quad k = 1, 2, 3.$$

For the free equation (2.11):

$$(40.24) \quad S_0(x) = -\psi(x) \cdot \psi(x), \quad S_k(x) = \frac{1}{m}i\nabla_k \psi(x) \cdot \psi(x), \quad k = 1, 2, 3.$$

41 Cauchy Problem for Maxwell Equations

We prove the existence of dynamics for inhomogeneous Maxwell equations and construct an integral representation for the solutions.

Let us consider the Cauchy problem for the Maxwell equations (4.1) with the initial conditions

$$(41.1) \quad \mathbf{E}|_{t=0} = \mathbf{E}_0(\mathbf{x}), \quad \mathbf{B}|_{t=0} = \mathbf{B}_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

We assume $(\mathbf{E}_0(\mathbf{x}), \mathbf{B}_0(\mathbf{x})) \in L^2 \oplus L^2$, where $L^2 = L^2(\mathbb{R}^3) \otimes \mathbb{R}^3$, $\rho(t, \mathbf{x}) \in C(\mathbb{R}, L^2(\mathbb{R}^3))$, $\mathbf{j}(t, \mathbf{x}) \in C(\mathbb{R}, L^2)$ and also $(\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x})) \in C(\mathbb{R}, L^2 \oplus L^2)$. Then the system (4.1) leads to the identities

$$(41.2) \quad \operatorname{div} \mathbf{E}_0(\mathbf{x}) = 4\pi\rho(0, \mathbf{x}) \quad \operatorname{div} \mathbf{B}_0(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^3,$$

which are necessary constraints for the existence of solutions to the overdetermined system (4.1).

Theorem 41.1 *Let $\mathbf{E}_0(x), \mathbf{B}_0(x)$ and $\rho(t, \mathbf{x}), \mathbf{j}(t, \mathbf{x})$ satisfy all conditions mentioned above and the constraints (4.2) and (41.2). Then*

- i) The Cauchy problem (4.1), (41.1) has a unique solution $(\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x})) \in C(\mathbb{R}, L^2 \oplus L^2)$.*
- ii) Let $\mathbf{j}(t, \mathbf{x}) \equiv 0$. Then the energy is conserved:*

$$(41.3) \quad \int_{\mathbb{R}^3} [\mathbf{E}^2(t, \mathbf{x}) + \mathbf{B}^2(t, \mathbf{x})] d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}^3.$$

- iii) The convolution representation holds*

$$(41.4) \quad \begin{pmatrix} \mathbf{E}(t) \\ \mathbf{B}(t) \end{pmatrix} = \mathbf{M}(t) * \begin{pmatrix} \mathbf{E}_0 \\ \mathbf{B}_0 \end{pmatrix} + 4\pi \int_0^t \mathbf{G}(t-s) * \begin{pmatrix} \rho(s) \\ \frac{1}{c} \mathbf{j}(s) \end{pmatrix} ds, \quad t \in \mathbb{R},$$

where $\mathbf{E}(t) := \mathbf{E}(t, \cdot)$ etc, and $\mathbf{M}(t)$ resp. $\mathbf{G}(t)$ is 6×6 resp. 6×4 matrix-valued distribution concentrated on the sphere $|\mathbf{x}| = |t|$, for every fixed $t \in \mathbb{R}$:

$$(41.5) \quad \mathbf{M}(t)(\mathbf{x}) = 0, \quad \mathbf{G}(t, \mathbf{x}) = 0, \quad \text{for } |\mathbf{x}| \neq |t|.$$

Proof *ad i)* We introduce the complex field $\mathbf{C}(t, \mathbf{x}) = \mathbf{E}(t, \mathbf{x}) + i\mathbf{B}(t, \mathbf{x})$ and rewrite (4.1) as

$$(41.6) \quad \frac{1}{c} \dot{\mathbf{C}}(t, \mathbf{x}) = -i \operatorname{rot} \mathbf{C}(t, \mathbf{x}) - \frac{4\pi}{c} \mathbf{j}(t, \mathbf{x}), \quad \mathbf{C}|_{t=0} = \mathbf{C}_0(\mathbf{x}),$$

$$(41.7) \quad \operatorname{div} \mathbf{C}(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}),$$

where $\mathbf{C}_0(\mathbf{x}) = \mathbf{E}_0(\mathbf{x}) + i\mathbf{B}_0(\mathbf{x})$. Fourier transformation $\hat{\mathbf{C}}(\mathbf{k}, t) = \int \exp(i\mathbf{k} \cdot \mathbf{x}) \mathbf{C}(t, \mathbf{x}) d\mathbf{x}$ leads to the equations

$$(41.8) \quad \dot{\hat{\mathbf{C}}}(t, \mathbf{k}) = c\hat{m}(\mathbf{k})\hat{\mathbf{C}}(t, \mathbf{k}) - 4\pi\hat{\mathbf{j}}(t, \mathbf{k}), \quad \hat{\mathbf{C}}|_{t=0} = \hat{\mathbf{C}}_0(\mathbf{k}),$$

$$(41.9) \quad -i\mathbf{k} \cdot \hat{\mathbf{C}}(\mathbf{k}, t) = 4\pi\hat{\rho}(\mathbf{k}, t),$$

where $\hat{m}(\mathbf{k})$ denotes the 3×3 skew-adjoint matrix of the operator $-\mathbf{k} \times$ in \mathbb{C}^3 . The solution $\hat{\mathbf{C}}(t, \mathbf{k})$ is defined uniquely from the first equation (41.8) of the overdetermined system (41.8), (41.9),

$$(41.10) \quad \hat{\mathbf{C}}(t, \mathbf{k}) = \exp(c\hat{m}(\mathbf{k})t)\hat{\mathbf{C}}_0(\mathbf{k}) - 4\pi \int_0^t \exp(c\hat{m}(\mathbf{k})(t-s))\hat{\mathbf{j}}(s, \mathbf{k}) ds, \quad \mathbf{k} \in \mathbb{R}^3.$$

We still have to show that (41.10) satisfies the constraint (41.9). Indeed, the Fourier transformed equations (41.2), (4.2) are

$$(41.11) \quad -i\mathbf{k} \cdot \hat{\mathbf{C}}_0(\mathbf{k}) = 4\pi\hat{\rho}(0, \mathbf{k}), \quad \mathbf{k} \in \mathbb{R}^3,$$

$$(41.12) \quad \dot{\hat{\rho}}(t, \mathbf{k}) - i\mathbf{k} \cdot \hat{\mathbf{j}}(t, \mathbf{k}) = 0, \quad \mathbf{k} \in \mathbb{R}^3, t \in \mathbb{R}.$$

With $S(t, \mathbf{k}) = 4\pi\hat{\rho}(t, \mathbf{k}) + i\mathbf{k} \cdot \hat{\mathbf{C}}(t, \mathbf{k})$ they imply by (41.8)

$$(41.13) \quad S(\mathbf{k}, 0) = 4\pi\hat{\rho}(\mathbf{k}, 0) + i\mathbf{k} \cdot \hat{\mathbf{C}}_0(\mathbf{k}) = 0, \quad \dot{S}(t, \mathbf{k}) = 4\pi\dot{\hat{\rho}}(t, \mathbf{k}) - 4\pi i\mathbf{k} \cdot \hat{\mathbf{j}}(t, \mathbf{k}) = 0, \quad \mathbf{k} \in \mathbb{R}^3,$$

since $\mathbf{k} \cdot \hat{\mathbf{m}}(\mathbf{k})\hat{\mathbf{C}}(t, \mathbf{k}) = 0$. Therefore, $S(t, \mathbf{k}) = 0$ which means (41.9). Since $\hat{\mathbf{m}}(\mathbf{k})$ is a skew-adjoint matrix, its exponent $\exp(c\hat{\mathbf{m}}(\mathbf{k})t)$ is unitary. Now i) follows from (41.10).

ad ii) The Parseval identity implies

$$(41.14) \quad \int_{\mathbb{R}^3} [\mathbf{E}^2(t, \mathbf{x}) + \mathbf{B}^2(t, \mathbf{x})] d\mathbf{x} = \int_{\mathbb{R}^3} |\mathbf{C}(t, \mathbf{x})|^2 d\mathbf{x} = (2\pi)^{-3} \int_{\mathbb{R}^3} |\hat{\mathbf{C}}(t, \mathbf{k})|^2 d\mathbf{k}.$$

Therefore, (41.3) follows from (41.10) since $\mathbf{j}(t, \mathbf{k}) \equiv 0$ and $\exp(c\hat{\mathbf{m}}(\mathbf{k})t)$ is a unitary matrix.

ad iii) We have to transform (41.10) back to position space in order to check (41.5). We have $\hat{m} = \hat{\mathbf{m}}(\mathbf{k}) = -\mathbf{k} \times$, $\hat{m}^2 = -\mathbf{k}^2 + |\mathbf{k} \times \mathbf{k}|$, $\hat{m}^3 = -|\mathbf{k}|^2 \hat{m}$, ... Hence,

$$\hat{m}^{2j+1} = (-1)^j |\mathbf{k}|^{2j} \hat{m} = (-1)^j \frac{\hat{m}}{|\mathbf{k}|} |\mathbf{k}|^{2j+1} \text{ for } j \geq 0,$$

$$\hat{m}^{2j} = \hat{m}^{2j-1} \hat{m} = (-1)^{j-1} |\mathbf{k}|^{2j-2} \hat{m}^2 = -(-1)^j \left(\frac{\hat{m}}{|\mathbf{k}|} \right)^2 |\mathbf{k}|^{2j}, \quad j \geq 1,$$

which yields by Euler's trick for the exponential

$$(41.15) \quad \begin{aligned} \exp(\hat{m}(\mathbf{k})t) &= \sum_0^\infty (\hat{m}t)^n / n! = \sum_0^\infty (\hat{m}t)^{2j} / (2j)! + \sum_0^\infty (\hat{m}t)^{2j+1} / (2j+1)! \\ &= 1 + \left(\frac{\hat{m}}{|\mathbf{k}|} \right)^2 (1 - \cos |\mathbf{k}|t) + \frac{\hat{m}}{|\mathbf{k}|} \sin |\mathbf{k}|t = \cos |\mathbf{k}|t + \hat{m} \frac{\sin |\mathbf{k}|t}{|\mathbf{k}|} + (1 - \cos |\mathbf{k}|t) \frac{|\mathbf{k} \times \mathbf{k}|}{|\mathbf{k}|^2}. \end{aligned}$$

Let us denote by $\hat{K}(t, \mathbf{k}) = \sin |\mathbf{k}|t / |\mathbf{k}|$, $\hat{\mathbf{m}}(t, \mathbf{k}) = \partial_t \hat{K}(t, \mathbf{k}) + \hat{m} \hat{K}(t, \mathbf{k})$, and $\hat{D}(t, \mathbf{k}) = 1 - \cos |\mathbf{k}|t$. Then we finally obtain,

$$(41.16) \quad \exp(c\hat{m}(\mathbf{k})t) = \hat{\mathbf{m}}(ct, \mathbf{k}) + |\mathbf{k} \times \frac{\hat{D}(ct, \mathbf{k})}{|\mathbf{k}|^2} < \mathbf{k}|.$$

Inserting this into (41.10) and using the constraints (41.11) and (41.12), we get

$$(41.17) \quad \hat{\mathbf{C}}(t, \mathbf{k}) = \hat{\mathbf{m}}(ct, \mathbf{k})\hat{\mathbf{C}}_0(\mathbf{k}) + 4\pi i |\mathbf{k} \times \frac{\hat{D}(ct, \mathbf{k})}{|\mathbf{k}|^2} \hat{\rho}(\mathbf{k}, 0)$$

$$(41.18) \quad -4\pi \int_0^t [\hat{\mathbf{m}}(c(t-s), \mathbf{k})\hat{\mathbf{j}}(s, \mathbf{k}) - i |\mathbf{k} \times \frac{\hat{D}(c(t-s), \mathbf{k})}{|\mathbf{k}|^2} \dot{\hat{\rho}}(s, \mathbf{k})] ds,$$

which through integration by parts becomes

$$(41.19) \quad \hat{\mathbf{C}}(t, \mathbf{k}) = \hat{\mathbf{m}}(ct, \mathbf{k})\hat{\mathbf{C}}_0(\mathbf{k}) - 4\pi c \int_0^t [\hat{\mathbf{m}}(c(t-s), \mathbf{k})\hat{\mathbf{j}}(s, \mathbf{k}) - i |\mathbf{k} \times \frac{\partial_t \hat{D}(c(t-s), \mathbf{k})}{|\mathbf{k}|^2} \hat{\rho}(s, \mathbf{k})] ds.$$

Using $\partial_t \hat{D}(t, \mathbf{k}) = |\mathbf{k}| \sin |\mathbf{k}|t = |\mathbf{k}|^2 \hat{K}(t, \mathbf{k})$, we get the inverse Fourier transforms,

$$(41.20) \quad \mathbf{m}(t, \mathbf{x}) := F_{\mathbf{k} \rightarrow \mathbf{x}}^{-1} \hat{\mathbf{m}}(t, \mathbf{k}) = \partial_t K(t, \mathbf{x}) - i \operatorname{rot} \circ K(t, \mathbf{x}),$$

$$(41.21) \quad \mathbf{g}(t, \mathbf{x}) := F_{\mathbf{k} \rightarrow \mathbf{x}}^{-1} \left(i|\mathbf{k}| \hat{K}(t, \mathbf{k}), -\hat{\mathbf{m}}(t, \mathbf{k}) \right) = (-\nabla K(t, \mathbf{x}), -\mathbf{m}(t, \mathbf{x})),$$

where $K(t, x)$ denotes the Kirchhoff kernel

$$(41.22) \quad K(t, \mathbf{x}) := F^{-1} \hat{K}(t, \mathbf{k}) = \frac{1}{4\pi t} \delta(|\mathbf{x}| - |t|).$$

With these notations, (41.19) implies (41.4) in the “complex” form

$$(41.23) \quad \mathbf{C}(t) = \mathbf{m}(ct) * \mathbf{C}_0 + 4\pi \int_0^t \mathbf{g}(c(t-s)) * \begin{pmatrix} c\rho(s) \\ \mathbf{j}(s) \end{pmatrix} ds \quad t \in \mathbb{R}.$$

Separating into real and imaginary parts, we obtain

$$(41.24) \quad \mathbf{E}(t, \mathbf{x}) = \mathbf{E}_{(r)}(t, \mathbf{x}) + \mathbf{E}_{(0)}(t, \mathbf{x}), \quad \mathbf{B}(t, \mathbf{x}) = \mathbf{B}_{(r)}(t, \mathbf{x}) + \mathbf{B}_{(0)}(t, \mathbf{x}),$$

where we denote

$$(41.25) \quad \begin{pmatrix} \mathbf{E}_{(0)}(t) \\ \mathbf{B}_{(0)}(t) \end{pmatrix} = \begin{pmatrix} \partial_t K(ct) & \operatorname{rot} \circ K(ct) \\ -\operatorname{rot} \circ K(ct) & \partial_t K(ct) \end{pmatrix} * \begin{pmatrix} \mathbf{E}_0 \\ \mathbf{B}_0 \end{pmatrix},$$

and

$$(41.26) \quad \begin{pmatrix} \mathbf{E}_{(r)}(t) \\ \mathbf{B}_{(r)}(t) \end{pmatrix} = 4\pi \int_0^t \begin{pmatrix} -\nabla K(c(t-s)) & -\partial_t K(c(t-s)) \\ 0 & \operatorname{rot} \circ K(c(t-s)) \end{pmatrix} * \begin{pmatrix} c\rho(s) \\ \mathbf{j}(s) \end{pmatrix} ds.$$

Here $K(t, \mathbf{x})$ coincides with the Kirchhoff kernel

$$(41.27) \quad K(t, \mathbf{x}) = F_{\mathbf{k} \rightarrow \mathbf{x}}^{-1} \hat{K}(t, \mathbf{k}) = \frac{1}{4\pi t} \delta(|\mathbf{x}| - |t|).$$

Now (41.4) and (41.5) follow immediately. \square

Remark The formula (41.26) coincides with standard Lienard-Wiechert representation of the “retarded” fields $\mathbf{E}_{(r)}(t, \mathbf{x})$ and $\mathbf{B}_{(r)}(t, \mathbf{x})$ through the Kirchhoff retarded potentials $\phi(t, \mathbf{x})$, $\mathbf{A}(t, \mathbf{x})$, [83]:

$$(41.28) \quad \begin{aligned} \mathbf{E}_{(r)}(t, \mathbf{x}) &= -\nabla \phi(t, \mathbf{x}) - \dot{\mathbf{A}}(t, \mathbf{x}), & \mathbf{B}_{(r)}(t, \mathbf{x}) &= \operatorname{rot} \mathbf{A}(t, \mathbf{x}), \\ \phi(t, \mathbf{x}) &= \int d^3y \frac{\Theta(\tau)}{4\pi|\mathbf{x} - \mathbf{y}|} \rho(\mathbf{y}, \tau), & \mathbf{A}(t, \mathbf{x}) &= \int d^3y \frac{\Theta(\tau)}{4\pi|\mathbf{x} - \mathbf{y}|} \mathbf{j}(\mathbf{y}, \tau), \end{aligned}$$

where $\tau = t - |\mathbf{x} - \mathbf{y}|$ is the *retarded time*, and $\Theta(\tau)$ is the Heaviside step function. We emphasize, that $(\mathbf{E}_{(r)}(t, \mathbf{x}), \mathbf{B}_{(r)}(t, \mathbf{x}))$ is not a solution to Maxwell equations (41.1) with prescribed $\rho(t, \mathbf{x})$ and $\mathbf{j}(t, \mathbf{x})$, since $\mathbf{E}_{(r)}|_{t=0} = 0$, and hence $\operatorname{div} \mathbf{E}(t, \mathbf{x}) = \rho(t, \mathbf{x})$ is not satisfied at $t = 0$. For the same reason, $(\mathbf{E}_{(0)}(t, \mathbf{x}), \mathbf{B}_{(0)}(t, \mathbf{x}))$ is not a solution to the Maxwell equations (41.1) with $\rho = 0, \mathbf{j} = 0$.

42 Lorentz Molecular Theory of Polarization and Magnetization

We analyze the *macroscopic* Maxwell field generated by charge and current distributions of a molecule at rest.

42.1 Constitutive Equations

The Maxwell equations (4.1) define the electromagnetic field generated by a given charge and current distribution. On the other hand, the Lorentz equations (4.21), (4.22) define the motion of charged particles in a given Maxwell field. However, the classical theory cannot explain the structure of matter, i.e., the stability of particles, the constitution of atoms and molecules, solid states etc. This is related to the fact that the coupled system (4.1), (4.21) (or (4.1), (4.21)) is not well defined for point particles. Hence, we miss the correct dynamical equation for matter. In particular, we need an additional hypothesis to get a satisfactory theory of matter in a Maxwell field.

Such a theory has been constructed by Lorentz [73] to justify the transition from a *microscopic* "electron theory" to the *macroscopic* Maxwell equations. First, the theory postulates that matter is a collection of *identical small cells* called *molecules*. Second, it is necessary to introduce an additional hypothesis concerning the molecular response to an external Maxwell field. The state of a neutral molecule is characterized by its *dipole moment and magnetic moment*. The parameters allow to describe the field generated by the molecule, at large distances from the molecule, with high precision. Hence, the parameters give a complete description of the molecular field for any macroscopic observation. Indeed, observations can be made only at distances which are much larger than the size of a molecule. Hence, it is sufficient to specify the influence of the external fields onto the parameters by the corresponding *constitutive equations*.

We start with an analysis of the distant Maxwell field generated by the charge and current distributions of a molecule at rest.

42.2 Stationary Molecular Fields in Dipole Approximation

Let us denote by $a > 0$ the size of the molecule and choose the origin 'in its center', i.e. assume that

$$(42.1) \quad \rho(t, \mathbf{y}) = 0, \quad \mathbf{j}(t, \mathbf{y}) = 0, \quad |\mathbf{x}| > a, \quad t \in \mathbb{R}.$$

Let us assume that

$$(42.2) \quad \rho(t, \cdot) \in L^1(\mathbb{R}^3), \quad \mathbf{j}(t, \cdot) \in L^1(\mathbb{R}^3) \otimes \mathbb{R}^3, \quad t \in \mathbb{R}.$$

Static fields First, let us consider the static case when the densities do not depend on time. Then Equations (4.6), (4.9) become the stationary Poisson equations and their solutions are the Coulomb potentials

$$(42.3) \quad \phi(\mathbf{x}) = \int \frac{\rho(\mathbf{y})d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|}, \quad \mathbf{A}(\mathbf{x}) = \frac{1}{c} \int \frac{\mathbf{j}(\mathbf{y})d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|}, \quad \mathbf{x} \in \mathbb{R}^3.$$

Let us expand $1/|\mathbf{x} - \mathbf{y}|$ in a Taylor series for small $|\mathbf{y}| \leq a$:

$$(42.4) \quad \frac{1}{|\mathbf{x} - \mathbf{y}|} = \frac{1}{\sqrt{\mathbf{x}^2 + \mathbf{y}^2 - 2\mathbf{y}\mathbf{x}}} = \frac{1}{|\mathbf{x}|} + \frac{\mathbf{y}\mathbf{x}}{|\mathbf{x}|^3} + \mathcal{O}\left(\frac{1}{|\mathbf{x}|^3}\right), \quad |\mathbf{x}| \rightarrow \infty.$$

Then (42.3) becomes,

$$(42.5) \quad \left\{ \begin{array}{l} \phi(\mathbf{x}) = \frac{Q}{|\mathbf{x}|} + \frac{\mathbf{p}\mathbf{x}}{|\mathbf{x}|^3} + \mathcal{O}\left(\frac{1}{|\mathbf{x}|^3}\right) \\ \mathbf{A}(\mathbf{x}) = \frac{\mathbf{J}}{c|\mathbf{x}|} + \frac{\mathcal{M}\mathbf{x}}{|\mathbf{x}|^3} + \mathcal{O}\left(\frac{1}{|\mathbf{x}|^3}\right) \end{array} \right. \quad |\mathbf{x}| \gg a,$$

where we denote

$$(42.6) \quad Q = \int \rho(\mathbf{y}) d\mathbf{y}, \quad \mathbf{p} = \int \mathbf{y} \rho(\mathbf{y}) d\mathbf{y},$$

$$\mathbf{J} = \int \mathbf{j}(\mathbf{y}) d\mathbf{y}, \quad \mathcal{M}_{kl} = \frac{1}{c} \int \mathbf{j}_k(\mathbf{y}) y_l d\mathbf{y}.$$

We will identify the molecular fields with the first two terms in the expansions (42.5), since $|\mathbf{x}|/a \gg 1$ in all *macroscopic observations*.

Let us note that the remainder in (42.4) is $\mathcal{O}(\mathbf{y}^2)$. Therefore, the first two terms in the expansions (42.5) correspond to the following *dipole approximations* for $\rho(\mathbf{y})$ and $\mathbf{j}(\mathbf{y})$:

$$(42.7) \quad \rho_d(\mathbf{y}) = Q\delta(\mathbf{y}) - \mathbf{p} \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}), \quad \mathbf{j}_d(\mathbf{y}) = \mathbf{J}\delta(\mathbf{y}) - c\mathcal{M}\nabla_{\mathbf{y}}\delta(\mathbf{y}).$$

42.3 Non-Stationary Fields in Multipole Approximations

It is easy to see that an asymptotic behavior of the type (42.5) holds for the *retarded* Kirchoff potentials (43.4) generated by non-stationary localized densities satisfying (42.1):

$$(42.8) \quad \phi(t, \mathbf{x}) = \int \frac{\rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y}, \quad A(t, \mathbf{x}) = \frac{1}{c} \int \frac{\mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y}, \quad (t, \mathbf{x}) \in \mathbb{R}^4.$$

For this purpose, let us continue the Taylor expansion (42.4) and obtain a complete expansion of the type (42.5), including all negative powers of $|\mathbf{x}|$:

$$(42.9) \quad \frac{1}{|\mathbf{x} - \mathbf{y}|} = \frac{1}{|\mathbf{x}|} + \frac{\mathbf{y}\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha| \geq 2} \frac{\mathbf{y}^\alpha P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}},$$

where we denote $\mathbf{n} := \mathbf{x}/|\mathbf{x}|$ and P_α is a polynomial. The expansion is a convergent series for $|\mathbf{x}| > a$. Hence, the retarded potentials may be expressed by the converging series

$$(42.10) \quad \left\{ \begin{array}{l} \phi(t, \mathbf{x}) = \frac{Q(t)}{|\mathbf{x}|} + \frac{\mathbf{p}(t)\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha| \geq 2} \frac{\phi_\alpha(t)P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}} \\ \mathbf{A}(t, \mathbf{x}) = \frac{\mathbf{J}(t)}{c|\mathbf{x}|} + \frac{\mathcal{M}(t)\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha| \geq 2} \frac{A_\alpha(t)P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}} \end{array} \right. \quad |\mathbf{x}| > a.$$

These expansions correspond to the following *multipole approximations* for $\rho(t, \mathbf{y})$ and $\mathbf{j}(t, \mathbf{y})$:

$$(42.11) \quad \left\{ \begin{array}{l} \rho_m(t, \mathbf{y}) = Q(t)\delta(\mathbf{y}) - \mathbf{p}(t) \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \rho_\beta(t)\nabla_{\mathbf{y}}^\beta\delta(\mathbf{y}), \\ \mathbf{j}_m(t, \mathbf{y}) = \mathbf{J}(t)\delta(\mathbf{y}) - c\mathcal{M}(t)\nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \mathbf{j}_\beta(t)\nabla_{\mathbf{y}}^\beta\delta(\mathbf{y}). \end{array} \right.$$

The coefficients are defined by

$$(42.12) \quad \left\{ \begin{array}{l} Q(t) = \int \rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}, \quad \mathbf{p}(t) = \int \mathbf{y} \rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}, \\ \mathbf{J}(t) = \int \mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}, \quad \mathcal{M}_{kl}(t) = \frac{1}{c} \int \mathbf{j}_k(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) y_l d\mathbf{y}, \end{array} \right.$$

$$(42.13) \quad \rho_\beta(t) = \int_{\mathbb{R}^3} \frac{(-\mathbf{y})^\beta}{\beta!} \rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}, \quad \mathbf{j}_\beta(t) = \int_{\mathbb{R}^3} \frac{(-\mathbf{y})^\beta}{\beta!} \mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y}) d\mathbf{y}.$$

Let us justify the convergence of the series (42.11) in the sense of distributions.

Definition 42.1 The space $\mathcal{H}_a(\mathbb{R}^3)$ consists of test functions $\psi(\mathbf{y})$ which are real analytic in the ball $B_a = \{\mathbf{y} \in \mathbb{R}^3 : |\mathbf{y}| < a\}$, and the Taylor series $\psi(\mathbf{y}) = \sum_{\beta} \frac{\mathbf{y}^{\beta}}{\beta!} \psi^{(\beta)}(0)$ converges uniformly in B_a .

Exercise 42.2 Check that the function $\psi_{\mathbf{x}}(\mathbf{y}) := 1/|\mathbf{x} - \mathbf{y}|$ belongs to the space $\mathcal{H}_a(\mathbb{R}^3)$ for $|\mathbf{x}| > a$.

Proposition 42.3 Let (42.1) hold and $\rho(t, \cdot), \mathbf{j}(t, \cdot) \in L^1(\mathbb{R}^3)$. Then the series (42.11) converges and coincides with $\rho(t, \mathbf{y})$ or $\mathbf{j}(t, \mathbf{y})$ in the following sense:

$$(42.14) \quad \begin{cases} \langle \rho(t, \mathbf{y}), \psi(\mathbf{y}) \rangle = \langle Q(t)\delta(\mathbf{y}) - \mathbf{p}(t) \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \rho_{\beta}(t) \nabla_{\mathbf{y}}^{\beta} \delta(\mathbf{y}), \psi(\mathbf{y}) \rangle \\ \langle \mathbf{j}(t, \mathbf{y}), \Psi(\mathbf{y}) \rangle = \langle \mathbf{J}(t)\delta(\mathbf{y}) - c\mathcal{M}(t)\nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \mathbf{j}_{\beta}(t) \nabla_{\mathbf{y}}^{\beta} \delta(\mathbf{y}), \Psi(\mathbf{y}) \rangle \end{cases}$$

for every test function from the space $\psi \in \mathcal{H}_a(\mathbb{R}^3)$ or $\Psi \in \mathcal{H}_a(\mathbb{R}^3) \otimes \mathbb{R}^3$, respectively.

Exercise 42.4 Prove Proposition 42.3. **Hint:** Substitute the Taylor expansion for ψ and Ψ into the LHS of (42.14) and use (42.1).

Corollary 42.5 The multipole approximations (42.11) satisfy the charge continuity equation

$$(42.15) \quad \dot{\rho}_m(t, \mathbf{y}) + \nabla_{\mathbf{y}} \cdot \mathbf{j}_m(t, \mathbf{y}) = 0, \quad (t, \mathbf{y}) \in \mathbb{R}^4.$$

Exercise 42.6 Prove the corollary. **Hint:** Use the identity $\langle \dot{\rho}_m(t, \mathbf{y}) + \nabla_{\mathbf{y}} \cdot \mathbf{j}_m(t, \mathbf{y}), \psi(t, \mathbf{y}) \rangle = \langle \dot{\rho}(t, \mathbf{y}) + \nabla_{\mathbf{y}} \cdot \mathbf{j}(t, \mathbf{y}), \psi(t, \mathbf{y}) \rangle$ for $\psi \in C_0^{\infty}(\mathbb{R}^4)$ such that $\psi(t, \cdot), \dot{\psi}(t, \cdot) \in \mathcal{H}_a(\mathbb{R}^3)$, $t \in \mathbb{R}$.

Substituting the series (42.11) into (42.15), we get

$$(42.16) \quad \dot{Q}(t)\delta(\mathbf{y}) - \dot{\mathbf{p}}(t) \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}) + \mathbf{J}(t) \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} C_{\beta}(t) \nabla_{\mathbf{y}}^{\beta} \delta(\mathbf{y}) = 0.$$

Therefore, we have

$$(42.17) \quad \dot{Q}(t) \equiv 0, \quad \mathbf{J}(t) \equiv \dot{\mathbf{p}}(t).$$

42.4 Magnetic Moment of a Molecule

Let us consider a molecule in a stationary state, i.e. $\rho(t, \mathbf{y}) \equiv \rho(\mathbf{y})$ and $\mathbf{j}(t, \mathbf{y}) \equiv \mathbf{j}(\mathbf{y})$. Then the multipole expansions (42.11) do not depend on time, i.e.,

$$(42.18) \quad \begin{cases} \rho_m(t, \mathbf{y}) \equiv \rho_m(\mathbf{y}) = Q\delta(\mathbf{y}) - \mathbf{p} \cdot \nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \rho_{\beta} \nabla_{\mathbf{y}}^{\beta} \delta(\mathbf{y}) \\ \mathbf{j}_m(t, \mathbf{y}) \equiv \mathbf{j}_m(\mathbf{y}) = \mathbf{J}\delta(\mathbf{y}) - c\mathcal{M}\nabla_{\mathbf{y}}\delta(\mathbf{y}) + \sum_{|\beta| \geq 2} \mathbf{j}_{\beta} \nabla_{\mathbf{y}}^{\beta} \delta(\mathbf{y}). \end{cases}$$

Proposition 42.7 Let the multipole charge-current densities (42.18) correspond to a stationary state of the molecule. Then the matrix \mathcal{M} is skewsymmetric, and

$$(42.19) \quad \mathcal{M}\nabla_{\mathbf{y}} = \mathbf{m} \times \nabla_{\mathbf{y}},$$

where $\mathbf{m} \in \mathbb{R}^3$ is the **magnetic moment** of the molecule, i.e., of the current density $\mathbf{j}(\mathbf{y})$ in the stationary state.

Proof Substituting (42.18) into (42.15), we get

$$(42.20) \quad \nabla_{\mathbf{y}} \cdot \mathbf{j}_m(\mathbf{y}) = 0, \quad \mathbf{y} \in \mathbb{R}^3.$$

Therefore, in particular $\mathbf{J} = 0$ and $\nabla_{\mathbf{y}} \cdot [\mathcal{M} \nabla_{\mathbf{y}} \delta(\mathbf{y})] = 0$. Then $\mathcal{M}_{kl} + \mathcal{M}_{lk} = 0$ and the vector $\mathbf{m} \in \mathbb{R}^3$ is defined by the following matrix identity:

$$(42.21) \quad \mathcal{M} = \begin{pmatrix} 0 & -\mathbf{m}_3 & \mathbf{m}_2 \\ \mathbf{m}_3 & 0 & -\mathbf{m}_1 \\ -\mathbf{m}_2 & \mathbf{m}_1 & 0 \end{pmatrix}. \quad \blacksquare$$

Formulas (42.21) and (42.6) imply that

$$(42.22) \quad \mathbf{m} = \frac{1}{2c} \int \mathbf{y} \times \mathbf{j}(\mathbf{y}) d\mathbf{y}.$$

Remark 42.8 *The integral does not depend on the choice of the origin since $\mathbf{J} := \int \mathbf{j}(\mathbf{y}) d\mathbf{y} = 0$ by (42.20).*

Adiabatic Condition Let us assume that the molecular dynamics can be described as an *adiabatic evolution* of stationary states with corresponding dipole electric moment $\mathbf{p}(t)$ and magnetic moment $\mathbf{m}(t)$.

Finally, let us assume that the molecule is neutral, i.e., $Q = 0$. We will identify $\rho(t, \mathbf{x})$ and $\mathbf{j}(t, \mathbf{x})$ with the first two terms on the RHS of (42.18). Then, by (42.17) and (42.19), we get the dipole approximation

$$(42.23) \quad \rho(t, \mathbf{x}) \approx \rho_d(t, \mathbf{x}) := -\mathbf{p}(t) \cdot \nabla_{\mathbf{x}} \delta(\mathbf{x}), \quad \mathbf{j}(t, \mathbf{x}) \approx \mathbf{j}_d(t, \mathbf{x}) := \dot{\mathbf{p}}(t) \delta(\mathbf{x}) + c \nabla_{\mathbf{x}} \times \mathbf{m}(t) \delta(\mathbf{x}).$$

Let us stress that the approximations are sufficient for any macroscopic observation of the molecular fields at distances much larger than the size of the molecules.

42.5 Macroscopic Limit: Maxwell Equations in Matter

Macroscopic Limit The total charge-current densities $\rho(t, \mathbf{x}), \mathbf{j}(t, \mathbf{x})$ in matter are sums of contributions of all molecules concentrated at the points $\mathbf{x}_n \in \mathbb{R}^3$. In the dipole approximation, the densities are

$$(42.24) \quad \begin{cases} \rho_{\text{mol}}(t, \mathbf{x}) &= -\sum_n \mathbf{p}_n(t) \cdot \nabla_{\mathbf{x}} \delta(\mathbf{x} - \mathbf{x}_n) \\ \mathbf{j}_{\text{mol}}(t, \mathbf{x}) &= \sum_n [\dot{\mathbf{p}}_n(t) \delta(\mathbf{x} - \mathbf{x}_n) + c \nabla_{\mathbf{x}} \times \mathbf{m}_n(t) \delta(\mathbf{x} - \mathbf{x}_n)]. \end{cases}$$

Now let us consider the *macroscopic limit*, when the diameter a of a molecule converges to zero, and each singular density converges to a limit distribution. More precisely, let us assume that for every fixed t we have the following asymptotics in the sense of distributions of \mathbf{x} :

$$(42.25) \quad \left\{ \begin{array}{l} \sum_n \mathbf{p}_n(t) \delta(\mathbf{x} - \mathbf{x}_n) \rightarrow \mathbf{P}(t, \mathbf{x}) \\ \sum_n \dot{\mathbf{p}}_n(t) \delta(\mathbf{x} - \mathbf{x}_n) \rightarrow \dot{\mathbf{P}}(t, \mathbf{x}), \quad \sum_n \mathbf{m}_n(t) \delta(\mathbf{x} - \mathbf{x}_n) \rightarrow \mathbf{M}(t, \mathbf{x}) \end{array} \right\} \Bigg|_{a \rightarrow 0}.$$

Then by the continuity of the differentiation of distributions, we have in the limit $a \rightarrow 0$,

$$(42.26) \quad \begin{cases} \rho_{\text{mol}}(t, \mathbf{x}) \approx -\nabla_{\mathbf{x}} \cdot \mathbf{P}(t, \mathbf{x}) \\ \mathbf{j}_{\text{mol}}(t, \mathbf{x}) \approx \dot{\mathbf{P}}(t, \mathbf{x}) + c\nabla_{\mathbf{x}} \times \mathbf{M}(t, \mathbf{x}) \end{cases}.$$

Definition 42.9 *i) The vector functions $\mathbf{P}(t, \mathbf{x})$ and $\mathbf{M}(t, \mathbf{x})$ are called the **electric polarization** and **magnetization** of the molecules at point \mathbf{x} and time t , respectively.*

*ii) $\mathbf{D}(t, \mathbf{x}) = \mathbf{E}(t, \mathbf{x}) + 4\pi\mathbf{P}(t, \mathbf{x})$ is called the **dielectric displacement** and $\mathbf{H}(t, \mathbf{x}) = \mathbf{B}(t, \mathbf{x}) - 4\pi\mathbf{M}(t, \mathbf{x})$ is called the **magnetic field intensity**.*

Let us separate the macroscopic and molecular charge and current densities:

$$(42.27) \quad \rho(t, \mathbf{x}) = \rho_{\text{mac}}(t, \mathbf{x}) + \rho_{\text{mol}}(t, \mathbf{x}), \quad \mathbf{j}(t, \mathbf{x}) = \mathbf{j}_{\text{mac}}(t, \mathbf{x}) + \mathbf{j}_{\text{mol}}(t, \mathbf{x}),$$

where the molecular densities are identified with the macroscopic limits (42.26). Let us substitute the expressions (42.27), (42.26) into the Maxwell equations (4.1) and express the fields E, B in terms of D, P, H, M . Then we obtain the *Maxwell equations in matter*:

$$(42.28) \quad \left\{ \begin{array}{ll} \operatorname{div} \mathbf{D}(t, \mathbf{x}) = 4\pi\rho_{\text{mac}}(t, \mathbf{x}), & \operatorname{rot} \mathbf{E}(t, \mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0, & \operatorname{rot} \mathbf{H}(t, \mathbf{x}) = \frac{1}{c}\dot{\mathbf{D}}(t, \mathbf{x}) + \frac{4\pi}{c}\mathbf{j}_{\text{mac}}(t, \mathbf{x}) \end{array} \right. \Bigg|_{(t, \mathbf{x}) \in \mathbb{R}^4}.$$

Constitutive Equations Equations (42.28) contain two additional unknown vector fields \mathbf{D}, \mathbf{H} . Therefore, we need two additional vector equations. For **isotropic** materials they are the **constitutive equations**

$$(42.29) \quad \mathbf{D}(t, \mathbf{x}) = \varepsilon\mathbf{E}(t, \mathbf{x}), \quad \mathbf{B}(t, \mathbf{x}) = \mu\mathbf{H}(t, \mathbf{x}),$$

where ε is called the **permittivity** and μ is called the **permeability** of matter. The **constitutive equations** are equivalent to

$$(42.30) \quad \mathbf{P}(t, \mathbf{x}) = \chi_e\mathbf{E}(t, \mathbf{x}), \quad \mathbf{M}(t, \mathbf{x}) = \chi_m\mathbf{H}(t, \mathbf{x}),$$

where χ_e is called the **electric susceptibility** and χ_m is called the **magnetic susceptibility** of matter. The following relations follow from Definition 42.9 ii):

$$(42.31) \quad \varepsilon = 1 + 4\pi\chi_e, \quad \mu = 1 + 4\pi\chi_m.$$

43 Long-Time Asymptotics and Scattering

The *retarded potentials* are particular solutions of the wave and Maxwell equations. Here we want to explain the outstanding role of these potentials.

43.1 Retarded Potentials

Formula (41.26) may be rewritten as

$$(43.1) \quad \mathbf{E}_{(r)}(t, \mathbf{x}) = -\nabla\phi_{(r)}(t, \mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}_{(r)}(t, \mathbf{x}), \quad \mathbf{B}_{(r)}(t, \mathbf{x}) = \text{rot } \mathbf{A}_{(r)}(t, \mathbf{x}),$$

where the potentials are given by

$$(43.2) \quad \phi_{(r)}(t, \mathbf{x}) = \int \frac{\Theta(t_{ret})}{|\mathbf{x} - \mathbf{y}|} \rho(t_{ret}, \mathbf{y}) d\mathbf{y}, \quad \mathbf{A}_{(r)}(t, \mathbf{x}) = \frac{1}{c} \int \frac{\Theta(t_{ret})}{|\mathbf{x} - \mathbf{y}|} \mathbf{j}(t_{ret}, \mathbf{y}) d\mathbf{y},$$

and $t_{ret} = t - |\mathbf{x} - \mathbf{y}|/c$. Let us assume that the charge and current densities are continuous and localized in space,

$$(43.3) \quad \rho(t, \mathbf{x}) = 0, \quad \mathbf{j}(t, \mathbf{x}) = 0, \quad |\mathbf{x}| > R, \quad t \in \mathbb{R}.$$

Then (43.2) for large $t > 0$ become the standard *retarded potentials* [53],

$$(43.4) \quad \left\{ \begin{array}{l} \phi_{(r)}(t, \mathbf{x}) = \phi_{ret}(t, \mathbf{x}) := \int \frac{\rho(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \\ \mathbf{A}_{(r)}(t, \mathbf{x}) = \mathbf{A}_{ret}(t, \mathbf{x}) := \frac{1}{c} \int \frac{\mathbf{j}(t - |\mathbf{x} - \mathbf{y}|/c, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y} \end{array} \right. \quad t > R + |\mathbf{x}|.$$

Further, the fields (43.1) become the *retarded fields*

$$(43.5) \quad \left\{ \begin{array}{l} \mathbf{E}_{(r)}(t, \mathbf{x}) = -\nabla\phi_{ret}(t, \mathbf{x}) - \dot{\mathbf{A}}_{ret}(t, \mathbf{x}), \\ \mathbf{B}_{(r)}(t, \mathbf{x}) = \text{rot } \mathbf{A}_{ret}(t, \mathbf{x}), \end{array} \right. \quad t > R + |\mathbf{x}|.$$

The distinguished role of the particular retarded solutions (43.4) to the wave equations (4.6), (4.9) is justified in *scattering theory*. In general, the solutions to the wave equations (4.6), (4.9) are defined uniquely by the initial conditions at time zero:

$$(43.6) \quad \phi|_{t=0} = \phi_0(\mathbf{x}), \quad \dot{\phi}|_{t=0} = \pi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

$$(43.7) \quad \mathbf{A}|_{t=0} = \mathbf{A}_0(\mathbf{x}), \quad \dot{\mathbf{A}}|_{t=0} = \Pi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3.$$

However, the **asymptotic behavior** of the solutions for $t \rightarrow +\infty$ and any **fixed** point \mathbf{x} does not depend on the initial data $\phi_0, \pi_0, \mathbf{A}_0, \Pi_0$ and coincides with the retarded potentials (43.4).

Space-localized initial data For example let us consider the initial functions with compact supports.

Proposition 43.1 *Let (43.3) hold, and let the initial functions $\phi_0(x), \pi_0(x), \mathbf{A}_0(x), \Pi_0(x)$ be continuous and localized in space,*

$$(43.8) \quad \phi_0(\mathbf{x}) = \pi_0(\mathbf{x}) = 0, \quad \mathbf{A}_0(\mathbf{x}) = \Pi_0(\mathbf{x}) = 0, \quad |\mathbf{x}| > R.$$

Then for large time the solutions to the Cauchy problems (4.6), (43.6) and (4.9), (43.7) coincide with the retarded potentials (43.4):

$$(43.9) \quad \phi(t, \mathbf{x}) = \phi_{ret}(t, \mathbf{x}), \quad \mathbf{A}(t, \mathbf{x}) = \mathbf{A}_{ret}(t, \mathbf{x}), \quad t > R + |\mathbf{x}|.$$

Proof Let us prove the proposition for the scalar potential $\phi(t, \mathbf{x})$. The Kirchhoff formula for the solution reads

$$(43.10) \quad \phi(t, \mathbf{x}) = \frac{1}{4\pi t} \int_{S_t(\mathbf{x})} \pi_0(\mathbf{y}) dS(\mathbf{y}) + \partial_t \left(\frac{1}{4\pi t} \int_{S_t(\mathbf{x})} \phi_0(\mathbf{y}) dS(\mathbf{y}) \right) + \phi_{ret}(t, \mathbf{x}), \quad t > 0, \quad \mathbf{x} \in \mathbb{R}^3,$$

where $S_t(\mathbf{x})$ denotes the sphere $\{\mathbf{y} \in \mathbb{R}^3 : |\mathbf{x} - \mathbf{y}| = t\}$ and $dS(\mathbf{y})$ is the Lebesgue measure on the sphere. Now (43.9) follows from (43.8). \blacksquare

A similar theorem holds for the Maxwell equations.

Theorem 43.2 *Let (43.3) and the conditions of Theorem 41.1 hold, and*

$$(43.11) \quad \mathbf{E}_0(\mathbf{x}) = \mathbf{B}_0(\mathbf{x}) = 0, \quad |\mathbf{x}| > R.$$

Then

$$(43.12) \quad \mathbf{E}(t, \mathbf{x}) = \mathbf{E}_{ret}(t, \mathbf{x}), \quad \mathbf{B}(t, \mathbf{x}) = \mathbf{B}_{ret}(t, \mathbf{x}) \quad t > R + |\mathbf{x}|.$$

Proof (43.12) follows from (41.24) and (43.5) since $\mathbf{E}_{(0)}, \mathbf{B}_{(0)}$ vanish for $t > R + |\mathbf{x}|$ by (43.11), (41.25) and (41.5). \blacksquare

Finite energy initial data For the Maxwell equations with $\mathbf{j} = 0$ the energy is conserved, (41.3). For the wave equations (4.6) and (4.9) with $\rho = 0$ and $\mathbf{j} = 0$ the energy conservations read (see (40.9)),

$$(43.13) \quad \int_{\mathbb{R}^3} \left[\frac{1}{c^2} |\dot{\phi}(t, \mathbf{x})|^2 + |\nabla \phi(t, \mathbf{x})|^2 \right] d\mathbf{x} = \text{const}, \quad \int_{\mathbb{R}^3} \left[\frac{1}{c^2} |\dot{\mathbf{A}}(t, \mathbf{x})|^2 + |\nabla \mathbf{A}(t, \mathbf{x})|^2 \right] d\mathbf{x} = \text{const}.$$

For solutions with finite initial energy, (43.9) and (43.12) hold in the local energy semi-norms in the limit $t \rightarrow +\infty$.

Theorem 43.3 *Let (43.3) hold and, further, $\phi(t, \mathbf{x})$, $\mathbf{A}(t, \mathbf{x})$, and $\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x})$ are finite energy solutions to the wave equations (4.6) and (4.9), or to the Maxwell Eqns (4.1), respectively. Then for any $R > 0$,*

$$(43.14) \quad \left\{ \begin{array}{l} \int_{|\mathbf{x}| < R} [|\dot{\phi}(t, \mathbf{x}) - \dot{\phi}_{ret}(t, \mathbf{x})|^2 + |\nabla \phi(t, \mathbf{x}) - \nabla \phi_{ret}(t, \mathbf{x})|^2] d\mathbf{x} \rightarrow 0 \\ \int_{|\mathbf{x}| < R} [|\dot{\mathbf{A}}(t, \mathbf{x}) - \dot{\mathbf{A}}_{ret}(t, \mathbf{x})|^2 + |\nabla \mathbf{A}(t, \mathbf{x}) - \nabla \mathbf{A}_{ret}(t, \mathbf{x})|^2] d\mathbf{x} \rightarrow 0 \\ \int_{|\mathbf{x}| < R} [|\mathbf{E}(t, \mathbf{x}) - \mathbf{E}_{ret}(t, \mathbf{x})|^2 + |\mathbf{B}(t, \mathbf{x}) - \mathbf{B}_{ret}(t, \mathbf{x})|^2] d\mathbf{x} \rightarrow 0 \end{array} \right. \quad t \rightarrow +\infty.$$

Proof Let us split the initial functions into two components: a first, space-localized component similar to (43.8), (43.11), and the rest. For the first component we construct a solution to the non-homogeneous equations, and for the rest we construct a solution to the homogeneous equations. Then, for the first component the convergence (43.14) follow from (43.9) and (43.12). The contribution of the second component is uniformly small in time by energy conservation. \blacksquare

43.2 Limiting Amplitude Principle in Scattering Problems

Time-periodic source and spectral problem Let us consider a wave problem with an external periodic source

$$(43.15) \quad (\partial_t^2 - \Delta)\phi(t, \mathbf{x}) = b(\mathbf{x})e^{-i\nu t},$$

where $\nu \in \mathbb{R}$ and the amplitude $b(\mathbf{x})$ decays rapidly as $|\mathbf{x}| \rightarrow \infty$ (see for example (12.9)). We assume $c = 1$ for simplicity. Similar problems arise in scattering problems for the Schrödinger equation

$$(43.16) \quad i\hbar\dot{\psi}(t, \mathbf{x}) = H\psi(t, \mathbf{x}) + b(\mathbf{x})e^{-i\nu t}$$

(see for example (13.2), (13.3)). Here H is the Schrödinger operator corresponding to a *static* Maxwell field with the potentials $\phi(t, \mathbf{x}) \equiv \phi(\mathbf{x})$ and $\mathbf{A}(t, \mathbf{x}) \equiv \mathbf{A}(\mathbf{x})$ (see (2.2)),

$$(43.17) \quad H := \frac{1}{2m}(-i\hbar\nabla_{\mathbf{x}} - \frac{e}{c}\mathbf{A}(\mathbf{x}))^2 + e\phi(\mathbf{x}).$$

Let us write Equations (43.15), (43.16) in a unified form,

$$(43.18) \quad i\dot{\Psi}(t, \mathbf{x}) = A\Psi(t, \mathbf{x}) + B(\mathbf{x})e^{-i\nu t}.$$

For the wave equation (43.15) we set $\Psi(t, \mathbf{x}) := (\phi(t, \mathbf{x}), \dot{\phi}(t, \mathbf{x}))$ and

$$(43.19) \quad A = i \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}, \quad B(\mathbf{x}) = i \begin{pmatrix} 0 \\ b(\mathbf{x}) \end{pmatrix}.$$

The corresponding *free equation* reads

$$(43.20) \quad i\dot{\Psi}(t, \mathbf{x}) = A\Psi(t, \mathbf{x}).$$

Let us denote by $U(t)$ the corresponding dynamical group if it exists:

$$(43.21) \quad \Psi(t) = U(t)\Psi(0),$$

where $\Psi(t) := \Psi(t, \cdot)$.

We consider *finite energy solutions* to Equations (43.18), (43.20). This means that $\Psi(t, \cdot) \in C(\mathbb{R}, E)$, where E is the corresponding Hilbert *phase space*. That is, $E := L^2(\mathbb{R}^3)$ for the Schrödinger equation (43.16), and $E := \{(\psi(\mathbf{x}), \pi(\mathbf{x})) : \nabla\psi(\mathbf{x}), \pi(\mathbf{x}) \in L^2(\mathbb{R}^3)\}$ is the Hilbert space with the norm

$$(43.22) \quad \|(\psi(\mathbf{x}), \pi(\mathbf{x}))\|_E^2 := \|\nabla\psi(\mathbf{x})\|^2 + \|\phi(\mathbf{x})\|^2,$$

where $\|\cdot\|$ stands for the norm in $L^2(\mathbb{R}^3)$.

Remark 43.4 $U(t)$, $t \in \mathbb{R}$, is a unitary operator in the corresponding Hilbert space by energy conservation.

Let us look for a solution of the type $\Psi_{\nu}(\mathbf{x})e^{-i\nu t}$ to Eq. (43.18). Substituting, we get the *Helmholtz* stationary equation,

$$(43.23) \quad (A - \nu)\Psi_{\nu}(\mathbf{x}) = -B(\mathbf{x}).$$

Definition 43.5 *i) The spectrum of Equation (43.18), $\text{spec } A$, is the set of all $\omega \in \mathbb{C}$ such that the operator $A + i\omega$ is not invertible in the corresponding Hilbert space E .*

ii) The resolvent of the operator A is defined by

$$(43.24) \quad R(\omega) := (A + i\omega)^{-1}, \quad \omega \in \mathbb{C} \setminus \text{spec } A.$$

Equation (43.23) admits a unique solution $\Psi_{\nu}(\mathbf{x}) \in E$ for every $B \in E$ if $\nu \notin \text{spec } A$:

$$(43.25) \quad \Psi_{\nu} = -R(\nu)B, \quad \nu \in \mathbb{C} \setminus \text{spec } A.$$

Exercise 43.6 Calculate $\text{spec } A$ for the free Schrödinger equation corresponding to (43.16) with $\hbar = 1$ and $m = 1$:

$$(43.26) \quad i\dot{\psi}(t, \mathbf{x}) = \Delta\psi(t, \mathbf{x}).$$

Hint: Use the Fourier transform $\hat{\psi}(\mathbf{k}) := \int e^{i\mathbf{k}\mathbf{x}}\psi(\mathbf{x})d\mathbf{x}$.

Solution: The operator $A = \Delta$ becomes the multiplication operator $\hat{A}(\mathbf{k}) = -\mathbf{k}^2$, and the phase space $\hat{E} = L^2(\mathbb{R}^3) \approx E$ by the Parseval theorem. Therefore, $\hat{R}(\omega, \mathbf{k})$ is multiplication by $-(\mathbf{k}^2 + \omega)^{-1}$. It is bounded in \hat{E} iff the function $(\mathbf{k}^2 + \omega)^{-1}$ is bounded in \mathbb{R}^3 . Hence, $\text{spec } A = \overline{\mathbb{R}}_- = \{\omega \in \mathbb{R} : \omega \leq 0\}$.

Exercise 43.7 Calculate the resolvent $R(\omega)$ corresponding to (43.26).

Hint: $R(\omega)$ is a convolution with a fundamental solution $\mathcal{E}_\omega(\mathbf{x})$ of the operator $\Delta - \omega$. Choose the fundamental solution from the space of tempered distributions.

Solution: The tempered fundamental solution is unique,

$$(43.27) \quad \mathcal{E}_\omega(\mathbf{x}) = \mathcal{E}_\omega^-(\mathbf{x}) := \frac{e^{-\sqrt{\omega}|\mathbf{x}|}}{4\pi|\mathbf{x}|}, \quad \omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_-,$$

where we choose $\text{Re } \sqrt{\omega} > 0$ for all $\omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_-$.

Remark 43.8 i) The distribution $\mathcal{E}_\omega(\mathbf{x})$ is tempered since its Fourier transform is a bounded function.

ii) The distribution $\mathcal{E}_\omega^+(\mathbf{x}) := \frac{e^{\sqrt{\omega}|\mathbf{x}|}}{4\pi|\mathbf{x}|}$ also is a fundamental solution, however, it is not tempered for all $\omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_-$.

iii) For $\omega < 0$, both fundamental solutions, $\mathcal{E}_\omega^\pm(\mathbf{x})$, are tempered.

Limiting amplitude principle The principle states (see [29]) the following long-time asymptotics for a general class of solutions to (43.18) with $\psi(0, \mathbf{x}) \in E$:

$$(43.28) \quad \Psi(t, \mathbf{x}) = \overline{\Psi}_\nu(\mathbf{x})e^{-i\nu t} + \sum_l C_l \Psi_l(\mathbf{x})e^{-i\omega_l t} + r(t, \mathbf{x}),$$

where $\overline{\Psi}_\nu(\mathbf{x})$ is a *limiting amplitude*, $\Psi_l(\mathbf{x})$ stand for the eigenfunctions of the discrete spectrum of the operator A , and $r(t, \mathbf{x}) \rightarrow 0$, $t \rightarrow \infty$, in an appropriate norm.

For the wave equation (43.15), the asymptotics (43.28) follows immediately from the Kirchhoff formula (43.10), if the initial functions decay rapidly. Let us assume, for example, that

$$(43.29) \quad |\nabla\phi(0, \mathbf{x})| + |\phi(0, \mathbf{x})| + |\pi(0, \mathbf{x})| = \mathcal{O}(e^{-\varepsilon|\mathbf{x}|}), \quad |\mathbf{x}| \rightarrow \infty,$$

with an $\varepsilon > 0$. Then the first and second terms on the RHS of (43.10) decay exponentially, like $e^{-\varepsilon t}$, while the last is just proportional to $e^{-i\nu t}$. Indeed, substituting $\rho(t, \mathbf{x}) = \frac{1}{4\pi}b(\mathbf{x})e^{-i\nu t}$ and $c = 1$ into (43.4), we get

$$(43.30) \quad \phi_{ret}(t, \mathbf{x}) := \int \frac{b(\mathbf{y})e^{-i\nu(t-|\mathbf{x}-\mathbf{y}|)}}{4\pi|\mathbf{x}-\mathbf{y}|}d\mathbf{y} = e^{-i\nu t}\overline{\phi}(\mathbf{x}), \quad \overline{\phi}(\mathbf{x}) = \int \frac{b(\mathbf{y})e^{i|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}d\mathbf{y}.$$

This gives (43.28) without the sum. This corresponds to the absence of the discrete spectrum for the wave equation with constant coefficients.

For the Schrödinger equation (43.16) the asymptotics (43.28) is proved now for a restricted class of potentials. For example, it can be deduced from the Kato and Jensen results [57] for

$$(43.31) \quad \phi(\mathbf{x}) = \mathcal{O}(|\mathbf{x}|^{-4-\varepsilon}), \quad \mathbf{A}(\mathbf{x}) \equiv 0,$$

where $\varepsilon > 0$. The asymptotics holds for the initial function $\psi(0, \mathbf{x})$ from the Agmon spaces, $H_\sigma = \{\psi(\mathbf{x}) : (1 + |\mathbf{x}|)^\sigma(1 - \Delta) \in L^2(\mathbb{R}^3)\}$ with $\sigma > 1/2$. Then the remainder $r(t, \mathbf{x}) \rightarrow 0$ in a dual space H_σ^* to H_σ .

Proposition 43.9 *i) For the Schrödinger equations (43.16), (43.17), with potentials satisfying (43.31), $\psi(0, \mathbf{x}) \in H_\sigma$, $\sigma > 1/2$, and $\nu \neq \omega_l$, $\nu \neq 0$, the asymptotics (43.28) holds and $\|r(t, \cdot)\|_{H_\sigma^*} \rightarrow 0$, $t \rightarrow \infty$.
ii) The limiting amplitude is given by (cf. (43.25))*

$$(43.32) \quad \bar{\Psi}_\nu = - \lim_{\varepsilon \rightarrow 0^+} R(\nu + i\varepsilon)B, \quad \nu \in \mathbb{R}, \quad B \in H_\sigma, \quad \sigma > 1/2,$$

where the limit holds in the space H_σ^* .

Remarks 43.10 *i) The limiting amplitude $\bar{\Psi}_\nu \in E := L^2(\mathbb{R}^3)$ if $\nu \notin \text{spec } A$. The results [57] imply that $\bar{\Psi}_\nu \in H_\sigma^*$ for $\nu \geq 0$, though $\nu \in \text{spec}_c A$ then, where $\text{spec}_c A$ stands for the **continuous spectrum** of the operator A . However generally $\bar{\Psi}_\nu \notin E$: this is related to a slow decay of $\bar{\Psi}_\nu(\mathbf{x})$ at infinity, $|\mathbf{x}| \rightarrow \infty$, which physically means a radiation of electrons to infinity.*

ii) The Coulomb potential $\phi(\mathbf{x}) = C|\mathbf{x}|^{-1}$ does not fit the bound (43.31).

iii) Formula (43.32) is called the “limiting absorption principle” since $R(\nu + i\varepsilon) = (A - \varepsilon + i\nu)^{-1}$ is the resolvent of the “damped” equation

$$(43.33) \quad i\dot{\Psi}(t, \mathbf{x}) = A\Psi(t, \mathbf{x}) - \varepsilon i\Psi(t, \mathbf{x}) + B(\mathbf{x})e^{-i\nu t},$$

where the term $-\varepsilon i\Psi(t, \mathbf{x})$ describes an absorption of energy.

Let us sketch a *formal* proof of the proposition using the results [57]. Namely, the Duhamel representation for the solutions to the inhomogeneous equation (43.18) reads,

$$(43.34) \quad \Psi(t) = U(t)\Psi(0) - i \int_0^t U(t-s)B e^{-i\nu s} ds,$$

where $U(t) := e^{-iAt}$ is the dynamical group of the corresponding free equation.

The results [57] imply that the first term on the RHS admits an asymptotics (43.28) with $\bar{\Psi}_\nu = 0$. This is a central achievement of the Kato-Jensen theory [57]: the contribution of the continuous spectrum to the solution $U(t)\Psi(0)$ of the homogeneous equation vanishes as $t \rightarrow \infty$.

It remains to analyze the integral term. It can be rewritten as

$$(43.35) \quad I(t) = -ie^{-i\nu t} \int_0^t U(s)B e^{i\nu s} ds.$$

Therefore, we have asymptotically,

$$(43.36) \quad I(t) = -ie^{-i\nu t} \int_0^t e^{-iAs} B e^{i\nu s} ds \sim e^{-i\nu t} \bar{\Psi}_\nu, \quad t \rightarrow \infty,$$

since

$$(43.37) \quad -i \int_0^\infty e^{-iAs} B e^{i\nu s} ds = -i \lim_{\varepsilon \rightarrow 0^+} \int_0^\infty e^{-iAs} B e^{i(\nu+i\varepsilon)s} ds = - \lim_{\varepsilon \rightarrow 0^+} (A - \nu - i\varepsilon)^{-1} B =: \bar{\Psi}_\nu,$$

where the last limit exists in the space H_σ^* by the results of [57] since $\nu \neq \omega_l$ and $\nu \neq 0$. ■

Remark 43.11 *Our arguments above demonstrate that*

*i) The term $\bar{\Psi}(\mathbf{x})e^{-i\nu t}$ on the RHS of (43.28) characterizes the **response** of a physical system to the external periodic source $B(\mathbf{x})e^{-i\nu t}$, and it does not depend on the initial data.*

ii) On the other hand, the sum over the discrete spectrum on the RHS depends only on the initial data and does not depend on the external source.

Hence, in an experimental observation of a scattering problem, the response, in principle, can be singled out if the experiment is concentrated on the field which is built by the external source.

Part VIII
Exercises

44 Exercises for Part I

44.1 Exercise 1: Main Lemma of the Calculus of Variations

Definition 44.1 $C[a, b]$ is a space of continuous functions on the interval $[a, b]$.

$$C_0[a, b] := \{h(x) \in C[a, b] : h(a) = h(b) = 0\}.$$

Lemma 44.2 Let $f(x) \in C[a, b]$ and $\int_0^T f(x)h(x) dx = 0$ for every $h(x) \in C_0[a, b]$. Then $f(x) \equiv 0$.

- a) Prove this lemma.
 b) Prove the analogous lemma for $h(x) \in C_0[a, b]$.

Proof of a) Let there exist a point x_0 such that $f(x_0) \neq 0$. Without loss of generality, we assume $f(x_0) = a > 0$. Since $f(x) \in C[a, b]$, then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - a| < \varepsilon$ for all $x \in \mathcal{O}_\delta(x_0) \equiv (x_0 - \delta, x_0 + \delta)$. We choose $\varepsilon = a/2$. Then $f(x) > a - \varepsilon = a/2$, $\forall x \in \mathcal{O}_\delta(x_0)$. Now we choose the function $h(x) \in C[a, b]$ such that $\text{supp } h(x) \in \mathcal{O}_\delta(x_0)$ and $h(x) \geq 0$. For example, $h(x) = 0$ for $|x - x_0| \geq \delta$ and $h(x) = \delta - |x - x_0|$ for $|x - x_0| \leq \delta$. Then

$$\int_0^T f(x)h(x) dx = \int_{x_0-\delta}^{x_0+\delta} f(x)h(x) dx > \frac{a}{2} \int_{x_0-\delta}^{x_0+\delta} h(x) dx > 0. \quad \square$$

The Proof of a) covers the case b), because the function $h(x)$ used in the proof $h(x) \in C_0[a, b]$.

44.2 Exercise 2: Euler-Lagrange Equations

- a) Given a functional $F : C^1[a, b] \rightarrow \mathbb{R}$,

$$F[y] := \int_a^b f(x, y(x), y'(x)) dx,$$

introduce a space

$$L[a, b] := \{y \in C^1[a, b], \quad y(a) = y_a, \quad y(b) = y_b\}.$$

Show that for a y which minimizes F , the Euler-Lagrange equation

$$f_y(x, y, y') - \frac{d}{dx} f_{y'}(x, y, y') = 0$$

holds. Here $f_y := \partial f / \partial y$.

Proof: Let $F(y_*) := \min_{y \in L[a, b]} F(y)$. Let $h(x) \in C_0^1[a, b] := \{h(x) \in C^1[a, b] : h(a) = h(b) = 0\}$.

Fix such a function $h(x)$. Introduce the function

$$S(\varepsilon) := F[y_* + \varepsilon h] = \int_a^b f(x, y_*(x) + \varepsilon h(x), y_*'(x) + \varepsilon h'(x)) dx, \quad \varepsilon \in \mathbb{R}.$$

Since $y_*(x) + \varepsilon h(x) \in L[a, b]$ for any ε , $S(\varepsilon) \geq S(0) = F(y_*)$. Hence, $\min S(\varepsilon) = S(0)$ and $S'(0) = 0$.

$$S'(\varepsilon) \Big|_{\varepsilon=0} = \int_a^b [f_y(x, y_*(x), y'_*(x))h(x) + f_{y'}(x, y_*(x), y'_*(x))h'(x)] dx = 0.$$

Integrating by parts we obtain

$$\int_a^b [f_y(x, y_*(x), y'_*(x)) - \frac{d}{dx} f_{y'}(x, y_*(x), y'_*(x))] h(x) dx = 0.$$

Exercise 1 implies $f_y(x, y_*(x), y'_*(x)) - \frac{d}{dx} f_{y'}(x, y_*(x), y'_*(x)) = 0$.

- b) **Geodesic line:** Use the Euler-Lagrange equations to calculate the minimum of the length functional in the plane

$$L[y] = \int_a^b \sqrt{1 + |y'(x)|^2} dx$$

with fixed boundary conditions $y(a) = y_a$, $y(b) = y_b$.

Solution: $f(x, y, y') := \sqrt{1 + |y'(x)|^2}$. Then $f_y = 0$. Substituting into the Euler-Lagrange equations we get

$$\frac{d}{dx} f_{y'}(x, y(x), y'(x)) = 0.$$

Hence, $f_{y'}(x, y(x), y'(x)) = \frac{y'}{\sqrt{1 + |y'(x)|^2}} = \text{const}$. Hence, $y' = \text{const}$, and $y(x) = a + bx$.

44.3 Exercise 3: Light Propagation in a Stratified Medium

Consider a medium where the index of refraction only depends on one coordinate y , $n(y) = 1/v(y)$ (here $v(y)$ is the speed of light.) The resulting system is effectively two-dimensional, i.e., the z coordinate may be suppressed.

- a) Use the **Fermat principle** (minimization of the flight time)

$$T[y] := T_{y(a), y(b)} := \int_a^b \frac{dl}{v(y)}$$

with fixed starting point $(x = a, y = y_a)$ and end point $(x = b, y = y_b)$ to prove the **Snellius law** of refraction

$$n(y) \sin \alpha(y) = \text{const}.$$

where $\alpha(y)$ is the angle between the light ray and the y -axis.

- b) Use the Snellius law to sketch the light rays of a wave guide with $n(y) = 2 + \cos y$.

- c) Use the Snellius law to sketch the light rays for ordinary refraction $n(y) = n_1$ for $y < 0$, and $n(y) = n_2$ for $y > 0$.

Proof of a) Since $dl = \sqrt{1 + (y'(x))^2}$ and $n(y) = 1/v(y)$, we rewrite $T[y]$ in the form

$$T[y] = \int_a^b n(y) \sqrt{1 + (y'(x))^2} dy.$$

Applying the Euler-Lagrange equations to $f(x, y, y') := n(y) \sqrt{1 + |y'(x)|^2}$, we get

$$n'(y) \sqrt{1 + |y'(x)|^2} - \frac{d}{dx} \frac{n(y)y'}{\sqrt{1 + |y'(x)|^2}} = 0.$$

Differentiating in x we get

$$\frac{n'(y)}{\sqrt{1 + |y'(x)|^2}} - n(y) \frac{y''}{(1 + |y'(x)|^2)^{3/2}} = 0.$$

This is equivalent to

$$\frac{d}{dx} \left[n(y) \cdot \frac{1}{\sqrt{1 + |y'(x)|^2}} \right] = 0.$$

Hence, $n(y) \cdot \frac{1}{\sqrt{1 + |y'(x)|^2}} = \text{const.}$ Since $\sin \alpha(y) = \frac{1}{\sqrt{1 + \text{ctg}^2 \alpha}} = \frac{1}{\sqrt{1 + |y'(x)|^2}}$, we get $n(y) \sin \alpha(y) = \text{const.}$ \square

45 Exercises for Part III

45.1 Exercise 4: The Kepler Problem in 5 Easy Steps

The Kepler problem is defined by the two-particle Lagrangian

$$L = \frac{m_1}{2} \dot{x}_1^2 + \frac{m_2}{2} \dot{x}_2^2 + \frac{\gamma m_1 m_2}{|x_1 - x_2|}.$$

Solve it by going through the following steps:

- Reduction to a one-particle problem (coordinate x) via separation of the center-of-mass coordinate X .
- Reduction to a two-dimensional problem by the observation that the plane spanned by $(x(0), \dot{x}(0))$ is invariant.
- Reduction to a one-dimensional problem (coordinate r) by the introduction of polar coordinates (r, φ) . Show that Kepler's third law holds, find the effective potential of the one-dimensional problem, and check for energy conservation.
- Investigate the orbits of the one-dimensional problem.
- Integrate the orbit equation to find the possible orbits. Show that, depending on the values of the integration constants (angular momentum, energy), the orbits are indeed ellipses, parabolae or hyperbolae.

Solution:

- a) Denote by $X(t)$ the center of mass, $X(t) := \frac{m_1 x_1(t) + m_2 x_2(t)}{M}$, where the total mass is $M := m_1 + m_2$, and $x := x_1 - x_2$. Then

$$\sum_{i=1}^2 m_i \dot{x}_i^2 = \frac{m_1 m_2 \dot{x}^2(t)}{M} + M \dot{X}^2.$$

Finally, we can rewrite the Lagrangian in the form

$$L = \frac{m_1 m_2 \dot{x}^2}{2M} + \frac{M \dot{X}^2}{2} + \frac{\gamma m_1 m_2}{|x_1 - x_2|} = L_1(x, \dot{x}, \dot{X}).$$

Applying the Euler-Lagrange equation to the coordinate X

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{X}} - \frac{\partial L}{\partial X} = 0,$$

we get $\frac{d}{dt} M \dot{X} = 0$. Hence,

$$(45.1) \quad \dot{X} = \text{const},$$

then $X(t) = at + b$, where a, b are some constants.

Remark 45.3 We can obtain Eqn (45.1) from the following arguments. Note that the potential

$$V(x_1, x_2) := -\frac{\gamma m_1 m_2}{|x_1 - x_2|}$$

depends only on the difference $x_1 - x_2$. Hence, V is invariant with respect to translations

$$T_s : (x_1, x_2) \rightarrow (x_1 + hs, x_2 + hs), \quad s \in \mathbb{R}, \quad h \in \mathbb{R}^3.$$

Hence, the total momentum (see Section 3.2) is conserved, $p := \sum_{i=1}^2 m_i \dot{x}_i = \text{const}$, and the center of mass moves uniformly, $X(t) = At + B$.

Finally, we get that it suffices to consider the Lagrangian

$$L_1 \equiv L_1(x, \dot{x}) := \frac{m_1 m_2 \dot{x}^2}{2M} + \frac{\gamma m_1 m_2}{|x|}.$$

- b) We can omit the constants $m_1 m_2$ and consider the equivalent Lagrangian

$$L \equiv L(x, \dot{x}) := \frac{\dot{x}^2}{2M} + \frac{\gamma}{|x|},$$

which gives the same trajectories. Denote by $M(t)$ the angular momentum $M(t) := x(t) \times p(t)$, where $p = \frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{M}$. Then,

$$(45.2) \quad M(t) = x(t) \times \frac{\dot{x}(t)}{M} = \text{const}, \quad t \in \mathbb{R}.$$

Indeed,

$$\dot{M}(t) = \dot{x}(t) \times p(t) + x(t) \times \dot{p}(t) = \dot{x}(t) \times \frac{\dot{x}(t)}{M} + x(t) \times \frac{\ddot{x}(t)}{M} = x(t) \times \frac{\ddot{x}(t)}{M}.$$

On the other hand, the Euler-Lagrange equation implies $\frac{\ddot{x}(t)}{M} + \frac{\gamma x}{|x|^3} = 0$. Hence, $\ddot{x} \parallel x$ and $\dot{M}(t) = x(t) \times \frac{\ddot{x}(t)}{M} = 0$. \square

Remark 45.4 We can obtain Eqn (45.2) from the following arguments. Note that the potential $V(x) := -\frac{\gamma}{|x|}$ depends only on $|x|$. Hence, $V(x)$ is invariant with respect to rotations (see section 2.3.2), and the angular momentum M is conserved (see Theorem 2.15).

Finally, Eq. (45.2) implies

$$x(t) \times \dot{x}(t) = x(0) \times \dot{x}(0) =: \bar{n}.$$

Hence, $x(t)$ belongs to the plane perpendicular to the vector \bar{n} . We can choose the coordinates such that $x_3 \parallel \bar{n}$. Then, $x(t) = (u(t), v(t), 0)$ and the Lagrangian for the variables $u(t), v(t)$ is

$$(45.3) \quad L(u, v, \dot{u}, \dot{v}) := \frac{\dot{u}^2 + \dot{v}^2}{2M} + \frac{\gamma}{\sqrt{u^2 + v^2}}.$$

c) We choose polar coordinates (r, φ) :

$$(45.4) \quad \begin{cases} u = r \cos \varphi, \\ v = r \sin \varphi. \end{cases}$$

Then $u^2 + v^2 = r^2$, $\dot{u}^2 + \dot{v}^2 = \dot{r}^2 + r^2 \dot{\varphi}^2$. We substitute into the Lagrangian and get

$$(45.5) \quad L(r, \varphi, \dot{r}, \dot{\varphi}) := \frac{\dot{r}^2 + r^2 \dot{\varphi}^2}{2M} + \frac{\gamma}{r}.$$

Applying the Euler-Lagrange equation to the coordinate φ , $\frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} = 0$, we get $\frac{d}{dt} \left[\frac{r^2 \dot{\varphi}}{M} \right] = 0$. Hence, we proved Kepler's third law

$$(45.6) \quad r^2 \dot{\varphi} = \text{const} =: I.$$

Remark 45.5 (45.6) is equivalent to (45.2). Namely, if we calculate the angular momentum $M(t)$ in polar coordinates, we get

$$M(t) = \frac{u\dot{v} - \dot{u}v}{M} = \frac{r^2 \dot{\varphi}}{M} = M(0).$$

Now applying the Euler-Lagrange equation to the coordinate r : $\frac{d}{dt} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0$, we get the reduced radial equation

$$\ddot{r} = r \dot{\varphi}^2 - \frac{\gamma M}{r^2}.$$

By (45.6) we rewrite it as follows:

$$\ddot{r} = \frac{I^2}{r^3} - \frac{\gamma M}{r^2} = -V'(r), \quad V(r) := \frac{I^2}{2r^2} - \frac{2\gamma M}{r}.$$

Multiplying this equation by \dot{r} and integrating, we obtain the reduced energy conservation

$$(45.7) \quad \dot{r}^2 + V(r) = E_0.$$

Remark 45.6 Since the Lagrangian (45.3) does not depend on t , the energy is conserved,

$$E(t) := \frac{\dot{u}^2 + \dot{v}^2}{2M} - \frac{\gamma}{\sqrt{u^2 + v^2}} = \text{const.}$$

In polar coordinates, taking into account (45.6), we get

$$E(t) = \frac{\dot{r}^2 + r^2\dot{\varphi}^2}{2M} - \frac{\gamma}{r} = \frac{\dot{r}^2}{2M} + \frac{I^2}{2Mr^2} - \frac{\gamma}{r} = \text{const.},$$

which coincides with (45.7).

d) Now we find $r(\varphi)$. Since $\frac{dr}{d\varphi} = \frac{\dot{r}}{\dot{\varphi}}$, we get by (45.7) and (45.2)

$$\frac{dr}{d\varphi} = \frac{\pm \sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E_0}}{\frac{I}{r^2}}.$$

Hence

$$\int \frac{I dr}{r^2 \sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E_0}} = \pm \int d\varphi.$$

We introduce the **Clerot substitution** $\rho := 1/r$. Then $dr = -\rho^{-2}d\rho$. Hence,

$$\int \frac{I dr}{r^2 \sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E_0}} = - \int \frac{I d\rho}{\sqrt{-I^2\rho^2 + 2\gamma M\rho + 2E_0}} = \int \frac{I d\rho}{\sqrt{D - I^2(\rho - B)^2}}.$$

Here $D := 2E_0 + (\gamma M)^2 I^{-2}$, $B := \gamma M I^{-2}$. Note that the constant D must be positive for any non-empty trajectory. Further,

$$\int \frac{I d\rho}{\sqrt{D - I^2(\rho - B)^2}} = \arcsin\left\{\frac{I(\rho - B)}{\sqrt{D}}\right\} = \pm\varphi + \varphi_0.$$

Finally, we get $\frac{I(\rho - B)}{\sqrt{D}} = \sin(\varphi + \varphi_0)$. Since $\rho := 1/r$, we get

$$(45.8) \quad \frac{A_1}{r} = A_2 + \sin(\varphi + \varphi_0),$$

where

$$(45.9) \quad A_1 := \frac{I}{\sqrt{D}},$$

$$(45.10) \quad A_2 := \frac{IB}{\sqrt{D}} = \frac{\gamma M}{\sqrt{2E_0 I^2 + (\gamma M)^2}}.$$

Note that r is bounded below by (45.8).

e) Let, for simplicity, $\varphi_0 = 0$. Then (45.8) becomes

$$(45.11) \quad A_2 r = A_1 - r \cos \varphi.$$

We remember that $r \sin \varphi = v$ and $r = \sqrt{u^2 + v^2}$, and now we re-write (45.11) in the form $(u^2 + v^2)A_2^2 = (A_1 - v)^2$. Hence,

$$u^2 + v^2 \left(1 - \frac{1}{A_2^2}\right) + 2\frac{A_1}{A_2^2}v = \frac{A_1^2}{A_2^2}.$$

Obviously, $A_1 \geq 0$. Let us assume that $A_1 > 0$. Then, if $A_2 > 1$ we have an ellipse, if $A_2 < 1$ we have a hyperbola, and if $A_2 = 1$ we have a parabola (A_2 is a positive constant). By (45.10) we get: if $E_0 < 0$ we have an ellipse, if $E_0 > 0$ we have a hyperbola, and if $E_0 = 0$ we have a parabola.

When $A_1 = 0$ (i.e. $I = 0$) then we have either $r = 0$, φ arbitrary (the point mass is in the origin) or $\sin \varphi = -A_2 = \text{const}$ (the point mass moves along a line through the origin).

45.2 Exercise 5: Bohr-Sommerfeld Quantization of the H Atom

The Hamiltonian for an electron with charge $-e$ and mass m in the Coulomb field of an infinitely heavy nucleus with charge e is

$$(45.12) \quad H = \frac{1}{2\mu} \left(p_r^2 + \frac{p_\varphi^2}{r^2} \right) - \frac{e^2}{r},$$

where the reduction to the effective two-dimensional problem (in polar coordinates (r, φ) , see (45.4)) has already been performed (cf (45.5)). Find the first two integrals of motion (conservation of angular momentum and energy). Assume that the energy E is less than zero to have periodic orbits (ellipses). Then use the Bohr-Sommerfeld quantization rules

$$(45.13) \quad \int p_r dr = kh, \quad \int p_\varphi d\varphi = lh$$

(here both integrals are over one period, k and l are integers, and h is Planck's constant) to derive the quantization of angular momentum and the Balmer formula for the energy levels of the electron. Use the principal quantum number $n = k + l$ in the latter case.

Solution: By (45.12) it follows that

$$(45.14) \quad \dot{\varphi} = H_{p_\varphi} = \frac{p_\varphi}{\mu r^2},$$

$$(45.15) \quad \dot{p}_\varphi = -H_\varphi = 0,$$

$$(45.16) \quad \dot{r} = H_{p_r} = \frac{p_r}{\mu},$$

$$(45.17) \quad \dot{p}_r = -H_r = \frac{p_\varphi^2}{\mu r^3} - \frac{e^2}{r^2},$$

By (45.15) we have

$$(45.18) \quad p_\varphi = M = \text{const}.$$

Then by (45.16) we get

$$(45.19) \quad \dot{p}_r = \frac{M^2}{\mu r^3} - \frac{e^2}{r^2}.$$

Since $\frac{dp_r}{dr} = \frac{\dot{p}_r}{\dot{r}}$, (45.16), (45.19) imply

$$(45.20) \quad \frac{dp_r}{dr} = \frac{\frac{M^2}{\mu r^3} - \frac{e^2}{r^2}}{\frac{p_r}{\mu}}, \quad \text{or} \quad \frac{p_r dp_r}{\mu} = \left(\frac{M^2}{\mu r^3} - \frac{e^2}{r^2} \right) dr.$$

Hence,

$$(45.21) \quad \frac{p_r^2}{2\mu} + \frac{M^2}{2\mu r^2} - \frac{e^2}{r} = \mathcal{E} = \text{const}.$$

Let $\mathcal{E} < 0$. Then (see Exercise 4) the orbits are ellipses. Hence,

$$(45.22) \quad p_r = \sqrt{2\mu\left(\frac{e^2}{r} - E\right) - \frac{M^2}{r^2}}.$$

where $E \equiv |\mathcal{E}| > 0$ was introduced for convenience.

Now we find the extrema of r . (45.16) implies that $\dot{r} = 0$ iff $p_r = 0$. By (45.22) we have

$$2\mu\left(\frac{e^2}{r} - E\right) - \frac{M^2}{r^2} = 0.$$

$$r^2 - \frac{e^2 r}{E} + \frac{M^2}{2\mu E} = 0.$$

$$r = \frac{e^2}{2E} \pm \sqrt{\frac{e^4}{4E^2} - \frac{M^2}{2\mu E}} = \frac{e^2}{2E} \pm \frac{1}{2E} \sqrt{e^4 - \frac{2M^2 E}{\mu}} =: r_{\pm}.$$

Hence, $r_{max} = r_+$ and $r_{min} = r_-$. Now we use (45.13). At first, (45.13) and (45.18) imply

$$(45.23) \quad \int p_r d\varphi = \int M d\varphi = 2\pi M = hl.$$

Hence, $M = hl$, $l \in \mathbb{Z}$. Then (45.22) implies

$$(45.24) \quad \begin{aligned} \int p_r dr &= 2 \int_{r_{min}}^{r_{max}} p_r dr = 2 \int_{r_{min}}^{r_{max}} \sqrt{2\mu\left(\frac{e^2}{r} - E\right) - \frac{M^2}{r^2}} dr \\ &= 2 \int_{r_{min}}^{r_{max}} \sqrt{-2\mu E r^2 + 2\mu e^2 r - M^2} \frac{dr}{r}. \end{aligned}$$

We use the formula

$$\begin{aligned} \int \sqrt{ax^2 + bx + c} \frac{dx}{x} &= \sqrt{ax^2 + bx + c} - \frac{b}{2\sqrt{-a}} \arcsin \frac{2ax + b}{b^2 - 4ac} \\ &\quad - \sqrt{-c} \arcsin \frac{bx + 2c}{x\sqrt{b^2 - 4ac}}, \quad a < 0, \quad c < 0. \end{aligned}$$

We apply this formula with $a = -2\mu E < 0$, $b = 2\mu e^2$, $c = -M^2 < 0$ to

$$(45.25) \quad \begin{aligned} \int p_r dr &= 2 \left[\sqrt{2\mu\left(\frac{e^2}{r} - E\right) - \frac{M^2}{r^2}} \Big|_{r_{min}}^{r_{max}} - \frac{2\mu e^2}{2\sqrt{2\mu E}} \arcsin \frac{-4\mu E r + 2\mu e^2}{\sqrt{4\mu^2 e^4 - 8M^2 \mu E}} \Big|_{r_{min}}^{r_{max}} \right. \\ &\quad \left. - |M| \arcsin \frac{2\mu e^2 r - 2M^2}{r\sqrt{4\mu^2 e^4 - 8M^2 \mu E}} \Big|_{r_{min}}^{r_{max}} \right] = 2 \left[0 + \pi \frac{\mu e^2}{\sqrt{2\mu E}} - |M| \pi \right] \\ &= 2\pi \left[\sqrt{\frac{\mu e^4}{2E}} - |M| \right]. \end{aligned}$$

Now (45.13) implies

$$2\pi \left[\sqrt{\frac{\mu e^4}{2E}} - |M| \right] = \hbar k, \quad k \in \mathbb{Z}.$$

Applying (45.23), we get

$$\sqrt{\frac{\mu e^4}{2E}} = \hbar(k + |l|).$$

Denoting $k + |l| = n$, we get

$$\frac{\mu e^4}{2E} = \hbar^2 n^2.$$

Hence, (cf. Theorem 8.1)

$$E = \frac{\mu e^4}{2\hbar^2 n^2} = \frac{R}{n^2}, \quad R := \frac{\mu e^4}{2\hbar^2}.$$

45.3 Exercise 6: Rutherford Scattering Formula

The orbit equation for a particle in a central potential $U(r)$ is

$$(45.26) \quad \varphi = \int_{r_0}^r \frac{M/r'^2 dr'}{\sqrt{2\mu(E - U(r')) - M^2/r'^2}},$$

where E is the energy and M is the angular momentum. For the Coulomb potential $U(r) = \alpha/r$ (with $\alpha \in \mathbb{R}$), and for $E > 0$ this is a scattering orbit (a hyperbola).

- a) Calculate the angle φ_0 , which is obtained from Eq. (1) by integrating from r_{\min} to ∞ . Express the total angle of deflection of the orbit for a particle starting at $r = \infty$ and going back to $r = \infty$ (i.e., the scattering angle χ) with the help of φ_0 .
- b) Re-express the constants of motion (E, M) by (E, b) , where b is the scattering parameter, i.e., the normal distance between the asymptote of the incident particle and the scattering center at $r = 0$.

Then, assume that a constant flux of particles (i.e., a constant number n of particles per area and time) with fixed energy and direction is approaching the scattering center. The number of particles per area and per time in a ring between b and $b + db$ is then $dN = 2\pi n b db$. If there is a one-to-one functional relation between b and the scattering angle χ , then dN is at the same time the number of particles per time that is scattered in an angle between χ and $\chi + d\chi$. Use this to calculate the differential cross section for Coulomb scattering (i.e., the Rutherford scattering formula)

$$d\sigma \equiv \frac{dN}{n} = 2\pi b(\chi) \left| \frac{db(\chi)}{d\chi} \right| d\chi.$$

Solution:

- a) For the Lagrangian $L = \frac{\mu \dot{x}^2}{2} + \frac{\alpha}{|x|}$ in polar coordinates, we get (45.26) (see Exercise 4). Let $E > 0$, then the orbits are hyperbolae. The scattering angle is $\chi = |\pi - 2\varphi_0|$, where

$$(45.27) \quad \varphi_0 := \int_{r_{\min}}^{\infty} \frac{M dr'}{r' \sqrt{2\mu E r'^2 - 2\mu \alpha r' - M^2}}.$$

We use the formula

$$\int \frac{dx}{x\sqrt{ax^2 + bx + c}} = \frac{1}{\sqrt{-c}} \arcsin \frac{bx + 2c}{x\sqrt{b^2 - 4ac}}, \quad c := -M^2 < 0.$$

Hence,

$$(45.28) \quad \varphi_0 := M \frac{1}{M} \arcsin \frac{-2\mu\alpha r' - 2M^2}{r' \sqrt{4\mu^2\alpha^2 + 8M^2\mu E}} \Big|_{r_{min}}^{\infty}.$$

Now we find $r = r_{min}$:

$$2\mu E r^2 - 2\mu\alpha r - M^2 = 0$$

Since $r > 0$, we find

$$r = \frac{\alpha}{E} + \sqrt{\frac{\alpha^2}{4E^2} + \frac{M^2}{2\mu E}} = \frac{\alpha}{E} + \frac{1}{2E} \sqrt{\alpha^2 + 2M^2 E/\mu}$$

$$(45.29) \quad \begin{aligned} \varphi_0 &:= \arcsin \frac{-\alpha - M^2/(\mu r)}{\sqrt{\alpha^2 + 2M^2 E/\mu}} \Big|_{r_{min}}^{\infty} = \arcsin \frac{-\alpha}{\sqrt{\alpha^2 + 2M^2 E/\mu}} - \arcsin(-1) \\ &= -\arcsin \frac{\alpha}{\sqrt{\alpha^2 + 2M^2 E/\mu}} + \pi/2. \end{aligned}$$

Hence,

$$(45.30) \quad \chi = |\pi - 2\varphi_0| = 2 \arcsin \frac{|\alpha|}{\sqrt{\alpha^2 + 2M^2 E/\mu}}.$$

b) Since $E = \mu v_{\infty}^2/2$, $v_{\infty} = \sqrt{2E/\mu}$. Hence, $M = \mu|\bar{r} \times \bar{v}| = \mu v_{\infty} b = \sqrt{2\mu E} b$. (45.30) implies

$$(45.31) \quad \chi = 2 \arcsin \frac{|\alpha|}{\sqrt{\alpha^2 + 4E^2 b^2}}.$$

Hence,

$$\frac{\alpha^2}{\sqrt{\alpha^2 + 4E^2 b^2}} = \sin^2 \frac{\chi}{2}.$$

Hence,

$$(45.32) \quad \begin{aligned} b^2 &= \frac{\alpha^2}{4E^2} \left[\frac{1}{\sin^2(\chi/2)} - 1 \right] = \frac{\alpha^2}{4E^2} \operatorname{ctg}^2 \frac{\chi}{2}, \\ b &= \frac{\alpha}{2E} \operatorname{ctg} \frac{\chi}{2}. \end{aligned}$$

At first, note that $dN = 2\pi n b db = 2\pi n b(\chi) \left| \frac{db(\chi)}{d\chi} \right| d\chi$ and $d\sigma \equiv \frac{dN}{n} = 2\pi b(\chi) \left| \frac{db(\chi)}{d\chi} \right| d\chi$. On the other side, then (45.32) implies

$$\frac{db}{d\chi} = \frac{\alpha}{2E} \frac{1}{2} \left(-\frac{1}{\sin^2(\chi/2)} \right).$$

Hence,

$$d\sigma = 2\pi \frac{\alpha}{2E} \operatorname{ctg} \frac{\chi}{2} \frac{\alpha}{2E} \frac{1}{2} \frac{1}{\sin^2(\chi/2)} d\chi = \pi \frac{\alpha^2}{4E^2} \frac{\cos \chi/2}{\sin^3 \chi/2} d\chi.$$

Hence, since $d\Omega = 2\pi \sin \chi d\chi = 4\pi \sin \chi/2 \cos \chi/2 d\chi$, we get for the differential scattering cross section

$$d\sigma = \frac{\alpha^2}{16E^2} \frac{1}{\sin^4 \chi/2} d\Omega.$$

45.4 Exercise 7: Energy Flow in Maxwell Field

Use the Maxwell equations in matter and the fact that $\int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j}$ is the change of kinetic energy per time of the charged matter contained in the volume Σ to show that $u = \frac{1}{8\pi}(\mathbf{E}^2 + \mathbf{B}^2)$ is the energy density of the electromagnetic field, and $\mathbf{S} \equiv \frac{c}{4\pi} \mathbf{E} \times \mathbf{B}$ is the density of energy flow.

Solution: Let us recall the Maxwell equations

$$\left\{ \begin{array}{l} \operatorname{div} \mathbf{E}(t, \mathbf{x}) = 4\pi\rho(t, \mathbf{x}), \quad \operatorname{rot} \mathbf{E}(t, \mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0, \quad \operatorname{rot} \mathbf{B}(t, \mathbf{x}) = \frac{1}{c}\dot{\mathbf{E}}(t, \mathbf{x}) + \frac{4\pi}{c}\mathbf{j}(t, \mathbf{x}), \end{array} \right. \quad (t, \mathbf{x}) \in \mathbb{R}^4,$$

where $\rho(t, \mathbf{x})$ and $\mathbf{j}(t, \mathbf{x})$ stand for the charge and current density, respectively. Subtracting $\mathbf{B} \cdot \nabla \times \mathbf{E}$ from $\mathbf{E} \cdot \nabla \times \mathbf{B}$ we obtain

$$(45.33) \quad \mathbf{E} \cdot \operatorname{rot} \mathbf{B} - \mathbf{B} \cdot \operatorname{rot} \mathbf{E} = \frac{1}{c} (\mathbf{E} \cdot \dot{\mathbf{E}} + \mathbf{B} \cdot \dot{\mathbf{B}}) + \frac{4\pi}{c} \mathbf{E} \cdot \mathbf{j}.$$

The l.h.s. may be rewritten like

$$(45.34) \quad \mathbf{E} \cdot \operatorname{rot} \mathbf{B} - \mathbf{B} \cdot \operatorname{rot} \mathbf{E} = (\nabla \cdot \mathbf{B}) \times \mathbf{E} - (\nabla \cdot \mathbf{E}) \times \mathbf{B} = -\nabla \cdot (\mathbf{E} \times \mathbf{B}),$$

where the brackets indicate the action of the derivative. We arrive at

$$(45.35) \quad \partial_t \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2) = -\mathbf{E} \cdot \mathbf{j} - \nabla \cdot \mathbf{S},$$

where

$$(45.36) \quad \mathbf{S} \equiv \frac{c}{4\pi} \mathbf{E} \times \mathbf{B}$$

is called Poynting vector. Integration over a region Σ of space leads to

$$(45.37) \quad \partial_t \frac{1}{8\pi} \int_{\Sigma} d^3x (\mathbf{E}^2 + \mathbf{B}^2) = - \int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j} - \int_{\partial\Sigma} da \mathbf{n} \cdot \mathbf{S}.$$

Next, we need the fact that $\int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j}$ is the change of kinetic energy per time unit of the matter described by \mathbf{j} in the electric field \mathbf{E} ,

$$(45.38) \quad \int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j} = \partial_t \mathcal{E}_{\text{kin}}.$$

This we will prove below. Now we integrate over the whole space in (45.37) and use that the fields go to zero for large distances to find

$$(45.39) \quad \partial_t \left(\int d^3x \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2) + \mathcal{E}_{\text{kin}} \right) = 0,$$

which is the conservation of the total energy. Therefore,

$$(45.40) \quad u = \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2)$$

is the energy density of the electromagnetic field, and the integral $\mathcal{E}_{\text{em}} = \int_{\Sigma} d^3x u$ is its energy in the volume Σ . For finite volume Σ we finally find

$$(45.41) \quad \partial_t (\mathcal{E}_{\text{em}} + \mathcal{E}_{\text{kin}}) = - \int_{\partial\Sigma} da \mathbf{n} \cdot \mathbf{S},$$

therefore $\int_{\partial\Sigma} da \mathbf{n} \cdot \mathbf{S}$ is the flow of electromagnetic energy through the surface $\partial\Sigma$, and \mathbf{S} is the density of the energy flow.

We still have to prove that $\int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j}$ is the power (change of kinetic energy per time unit) of the charged matter in the volume Σ in the electric field \mathbf{E} . Here we assume that the matter is composed of point particles and prove the assumption for a single particle, where $\mathbf{j} = e\mathbf{v}\delta(x - x(t))$. We find

$$(45.42) \quad \int_{\Sigma} d^3x \mathbf{E} \cdot \mathbf{j} = e\mathbf{v} \cdot \mathbf{E}(x),$$

if the particle is contained in the volume Σ . Further, we have in the non-relativistic case

$$(45.43) \quad \mathcal{E}_{\text{kin}} = \frac{m}{2} \mathbf{v}^2 \quad \Rightarrow \quad \partial_t \mathcal{E}_{\text{kin}} = m\mathbf{v} \cdot \dot{\mathbf{v}}$$

and the expression for the Lorentz force

$$(45.44) \quad m\dot{\mathbf{v}} = e\mathbf{E} + \frac{e}{c} \mathbf{v} \times \mathbf{B}.$$

Inserted into the expression for $\partial_t \mathcal{E}_{\text{kin}}$ this gives

$$(45.45) \quad \partial_t \mathcal{E}_{\text{kin}} = e\mathbf{v} \cdot \mathbf{E},$$

because the second (magnetic) term in the Lorentz force does not contribute. This is what we wanted to prove. The proof for relativistic matter particles is similar. ■

45.5 Exercise 8: Electromagnetic Plane Waves

Find the electric and magnetic fields of a plane electromagnetic wave by inserting the plane wave ansatz $\mathbf{E} = \mathbf{E}_0 \exp(i\mathbf{k} \cdot \mathbf{x} - i\omega t)$ into the free Maxwell equations. Specialize to $\mathbf{k} = k\hat{e}_3$, find the two linearly independent solutions and construct a circularly polarized wave as an example. Calculate the corresponding energy flow per time unit.

Solution: The Maxwell equations in vacuum are

$$(45.46) \quad \begin{aligned} \nabla \times \mathbf{B} &= \frac{1}{c} \dot{\mathbf{E}} \quad , \quad \nabla \times \mathbf{E} = -\frac{1}{c} \dot{\mathbf{B}} \\ \nabla \cdot \mathbf{B} &= 0 \quad , \quad \nabla \cdot \mathbf{E} = 0. \end{aligned}$$

From here the wave equation may be derived easily, e.g.,

$$(45.47) \quad \begin{aligned} \nabla \times (\nabla \times \mathbf{E}) &= -\frac{1}{c} \nabla \times \dot{\mathbf{B}}, \\ \nabla(\nabla \cdot \mathbf{E}) - \Delta \mathbf{E} &= -\frac{1}{c^2} \ddot{\mathbf{E}}, \\ \Delta \mathbf{E} &= \frac{1}{c^2} \ddot{\mathbf{E}}, \end{aligned}$$

and, in a similar fashion

$$(45.48) \quad \Delta \mathbf{B} = \frac{1}{c^2} \ddot{\mathbf{B}}.$$

Inserting the plane-wave ansatz

$$(45.49) \quad \mathbf{E} = \mathbf{E}_0 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t}$$

into the wave equation leads to the dispersion relation

$$(45.50) \quad \mathbf{k}^2 = \frac{\omega^2}{c^2}.$$

Further, the plane wave ansatz may be verified for the magnetic field with the help of the Maxwell equations. From

$$(45.51) \quad \nabla \times \mathbf{E} = i\mathbf{k} \times \mathbf{E}_0 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} = -\frac{1}{c} \dot{\mathbf{B}}$$

we find upon integration

$$(45.52) \quad \mathbf{B} = \frac{c}{\omega} \mathbf{k} \times \mathbf{E} + \mathbf{B}_0(\mathbf{x})$$

where \mathbf{B}_0 is a constant of integration which may depend on \mathbf{x} only. Further, we have

$$(45.53) \quad \nabla \times \mathbf{B} = \frac{1}{c} \mathbf{E},$$

$$(45.54) \quad -i\frac{c}{\omega} \mathbf{k}^2 \mathbf{E} + \nabla \times \mathbf{B}_0 = -i\frac{\omega}{c} \mathbf{E},$$

and, therefore, $\nabla \times \mathbf{B}_0 = 0$, which, together with $\nabla \cdot \mathbf{B}_0 = 0$, implies $\mathbf{B}_0 = \text{const.}$ If we assume that such a constant magnetic field in all space is absent, then

$$(45.55) \quad \mathbf{B} = \frac{\mathbf{k}}{|\mathbf{k}|} \times \mathbf{E} = \frac{\mathbf{k}}{|\mathbf{k}|} \times \mathbf{E}_0 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t}$$

follows. Obviously, $|\mathbf{E}| = |\mathbf{B}|$ and $\mathbf{E} \perp \mathbf{B}$ hold.

Inserting the plane wave ansatz into the Maxwell equations, we get

$$(45.56) \quad \begin{aligned} \mathbf{k} \times \mathbf{B} &= -\frac{\omega}{c} \mathbf{E} \quad , \quad \mathbf{k} \times \mathbf{E} = \frac{\omega}{c} \mathbf{B}, \\ \mathbf{k} \cdot \mathbf{B} &= 0 \quad , \quad \mathbf{k} \cdot \mathbf{E} = 0. \end{aligned}$$

Here we allow for a complex notation for the plane wave field for convenience, but one should keep in mind that only the real part of the field is physical. For a plane wave field (45.49) with $\mathbf{E}_0 = \mathbf{n}E_0 = \mathbf{n}|E_0|e^{i\phi}$ the physical field is

$$(45.57) \quad \mathbf{E}_{\text{ph}} = \text{Re} \left(\mathbf{E}_0 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} \right) = \mathbf{n}|E_0| \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \phi)$$

where \mathbf{n} is a real constant unit vector.

Now let us assume without loss of generality that $\mathbf{k} = k\hat{z}$. Then Equation (45.56) has two linearly independent solutions pointing into the x and y directions, and, consequently, the general solution is

$$(45.58) \quad \mathbf{E} = E_{0,x} \hat{e}_x e^{ikz - i\omega t} + E_{0,y} \hat{e}_y e^{ikz - i\omega t}$$

$$(45.59) \quad \mathbf{E}_{\text{ph}} = |E_{0,x}| \hat{e}_x \cos(kz - \omega t + \phi_1) + |E_{0,y}| \hat{e}_y \cos(kz - \omega t + \phi_2).$$

If $\phi_1 = \phi_2$, then the electromagnetic wave is linearly polarized. If we assume that $\phi_1 = \phi_2 = 0$ and focus on the x, y plane $z = 0$, then the electric field points into the $|E_{0,x}| \hat{e}_x + |E_{0,y}| \hat{e}_y$ direction at a time $t = 0$, is zero at a time $t = (\pi/2\omega)$, points into the minus $|E_{0,x}| \hat{e}_x + |E_{0,y}| \hat{e}_y$ direction at $t = (\pi/\omega)$, etc. Therefore, the electric field vector oscillates along a line parallel to the $|E_{0,x}| \hat{e}_x + |E_{0,y}| \hat{e}_y$ direction.

Circular polarization we have for $|E_{0,x}| = |E_{0,y}| \equiv E_0$ and $\phi_2 - \phi_1 = (\pi/2)$. For $\phi_1 = 0$, e.g., we have in the $z = 0$ plane an electric field vector with length E_0 pointing into the \hat{e}_x direction

at time $t = 0$, a vector with the same length pointing into the \hat{e}_y direction at time $t = (\pi/2\omega)$, etc. Therefore, the field vector rotates in a counter-clockwise direction with angular velocity ω without changing its length. We find for the physical fields in this case

$$(45.60) \quad \mathbf{E}_{\text{ph}} = E_0 (\hat{e}_x \cos(kz - \omega t) - \hat{e}_y \sin(kz - \omega t))$$

$$(45.61) \quad \mathbf{B} = \frac{c}{\omega} \mathbf{k} \times \mathbf{E} = \hat{e}_z \times \mathbf{E} = E_0 (\hat{e}_y \cos(kz - \omega t) + \hat{e}_x \sin(kz - \omega t))$$

and, therefore, for the energy flow density

$$(45.62) \quad \mathbf{S} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{B} = \frac{c}{4\pi} E_0^2 \hat{e}_z = \frac{c}{4\pi} E_0^2 \frac{\mathbf{k}}{|\mathbf{k}|}.$$

The energy flow density points into the direction of the wave vector \mathbf{k} .

Remark 45.7 For general polarizations the energy flow is still time dependent. E.g., for linear polarization $\mathbf{E} = E_0 \mathbf{n} \cos(kz - \omega t)$ (E_0 real, \mathbf{n} in the x, y plane), the energy flow density is

$$(45.63) \quad \mathbf{S} = \frac{c}{4\pi} E_0^2 \frac{\mathbf{k}}{|\mathbf{k}|} \cos^2(kz - \omega t).$$

Here, one is more interested in the average energy flow, where a time average for times $t \gg \omega^{-1}$ is performed,

$$(45.64) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \cos^2(kz - \omega t) dt = \frac{1}{2},$$

which leads to

$$(45.65) \quad \bar{\mathbf{S}} = \frac{c}{8\pi} E_0^2 \frac{\mathbf{k}}{|\mathbf{k}|}$$

(the factor of $(1/2)$ compared to (45.62) is due to the fact that in (45.65) we took into account only one of the two degrees of freedom).

45.6 Exercise 9: Hertzian Dipole Radiation

For a time-dependent dipole, $\rho(t, \mathbf{x}) = -\mathbf{p}(t) \cdot \nabla \delta^3(\mathbf{x})$, calculate the radiation field, i.e., the electric and magnetic fields in the radiation (far) zone. Calculate the corresponding energy flux and discuss its direction dependence. Hint: use the retarded potentials and the large distance expansion.

Solution: With the help of the current conservation equation $\dot{\rho} + \nabla \cdot \mathbf{j} = 0$, the corresponding dipole current density may be calculated,

$$(45.66) \quad \mathbf{j} = \dot{\mathbf{p}}(t) \delta^3(\mathbf{x}).$$

The retarded vector potential generated by this dipole is

$$(45.67) \quad \begin{aligned} \mathbf{A}(\mathbf{x}, t) &= \frac{1}{c} \int d^3x' \int dt' \frac{\mathbf{j}(x', t')}{|\mathbf{x} - \mathbf{x}'|} \delta(t' + \frac{|\mathbf{x} - \mathbf{x}'|}{c} - t) = \\ &= \frac{1}{c} \int dt' \dot{\mathbf{p}}(t') \frac{1}{|\mathbf{x}|} \delta(t' + \frac{|\mathbf{x}|}{c} - t) = \frac{1}{c} \frac{\dot{\mathbf{p}}(t - (r/c))}{r}, \end{aligned}$$

where $r = |\mathbf{x}|$. The scalar potential may be calculated in an easy way by employing the Lorentz gauge condition $\nabla \cdot \mathbf{A} + (1/c)\dot{\phi} = 0$,

$$\begin{aligned} \dot{\phi} &= -c\nabla \cdot \mathbf{A} = -\partial_t \nabla \cdot \frac{\mathbf{p}(t - (r/c))}{r}, \\ \phi &= -\partial_t \nabla \frac{\mathbf{p}(t - (r/c))}{r} = -\frac{1}{r} \nabla \cdot \mathbf{p}(t - (r/c)) - \mathbf{e}\mathbf{p} \cdot \nabla \frac{1}{r} = \\ (45.68) \quad &\frac{\mathbf{x} \cdot \dot{\mathbf{p}}(t - (r/c))}{cr^2} + \frac{\mathbf{x} \cdot \mathbf{p}}{r^3} \simeq \frac{\mathbf{x} \cdot \dot{\mathbf{p}}(t - (r/c))}{cr^2}, \end{aligned}$$

where the large distance approximation was used in the last step.

Now the magnetic field may be calculated,

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} = \frac{1}{c} \nabla \times \frac{\dot{\mathbf{p}}(t - (r/c))}{r} = \\ (45.69) \quad &\frac{1}{c} \dot{\mathbf{p}}(t - (r/c)) \times \frac{\mathbf{x}}{r^3} + \frac{1}{cr} \nabla(t - \frac{r}{c}) \times \dot{\mathbf{p}}(t - \frac{r}{c}) \simeq \frac{1}{c^2} \frac{\ddot{\mathbf{p}}(t - \frac{r}{c}) \times \mathbf{x}}{r^2}, \end{aligned}$$

and, in a similar fashion, the electric field

$$\begin{aligned} \mathbf{E} &= -\frac{1}{c} \dot{\mathbf{A}} - \nabla \phi \simeq -\frac{1}{c^2} \frac{\ddot{\mathbf{p}}}{r} - \nabla \frac{\mathbf{x} \cdot \dot{\mathbf{p}}(t - \frac{r}{c})}{cr^2} = \\ &-\frac{1}{c^2} \frac{\ddot{\mathbf{p}}}{r} - \frac{1}{cr^2} (\dot{\mathbf{p}} + x_j \nabla \dot{p}_j(t - \frac{r}{c})) - \frac{1}{c} \mathbf{x} \cdot \dot{\mathbf{p}} \nabla r^{-2} \simeq \\ (45.70) \quad &-\frac{1}{c^2} \frac{\ddot{\mathbf{p}}}{r} + \frac{\mathbf{x}(\mathbf{x} \cdot \ddot{\mathbf{p}})}{c^2 r^3} = \frac{1}{c^2 r^3} (\ddot{\mathbf{p}}(t - \frac{r}{c}) \times \mathbf{x}) \times \mathbf{x}, \end{aligned}$$

where for \mathbf{E} only the first and third terms in the second line contribute in the large distance limit. Observe that $\mathbf{x} \perp \mathbf{E} \perp \mathbf{B}$, i.e., this is a radiation field.

For the energy flow (Poynting vector) we find

$$\begin{aligned} \mathbf{S} &= \frac{c}{4\pi} \mathbf{E} \times \mathbf{B} = \frac{1}{4\pi c^3 r^5} \left((\ddot{\mathbf{p}} \cdot \mathbf{x}) \mathbf{x} - \ddot{\mathbf{p}} r^2 \right) \times (\ddot{\mathbf{p}} \times \mathbf{x}) = \\ (45.71) \quad &\frac{1}{4\pi c^3 r^5} \left(r^2 \ddot{\mathbf{p}}^2 - (\ddot{\mathbf{p}} \cdot \mathbf{x})^2 \right) \mathbf{x} \end{aligned}$$

If we assume that the dipole is oriented toward the z direction, $\mathbf{p} = (0, 0, p)$ then we find

$$(45.72) \quad \mathbf{S} = \frac{\sin^2 \theta}{4\pi c^3 r^2} \dot{p}^2 \mathbf{n},$$

where

$$(45.73) \quad \mathbf{n} = \frac{\mathbf{x}}{r}, \quad \cos \theta = \frac{\mathbf{n} \cdot \mathbf{p}}{|\mathbf{p}|}.$$

Therefore, \mathbf{S} behaves like r^{-2} for large distances (which it must due to energy conservation), and there is no radiation (energy flow) in the direction of the dipole (for $\theta = 0, \pi$).

45.7 Exercise 10: Polarizability in an External Electromagnetic Wave

In many instances, a classical, point-like electron in matter may be described approximately as a damped harmonic oscillator which oscillates around a fixed, positively charged center. In an external electromagnetic field the corresponding equation of motion for the position of the electron is

$$m(\ddot{\mathbf{x}}(t) + \gamma\dot{\mathbf{x}}(t) + \omega_0^2\mathbf{x}(t)) = e(\mathbf{E} + \frac{1}{c}\dot{\mathbf{x}} \times \mathbf{B}).$$

Here m is the mass and e the charge of the electron. γ is a damping constant and \mathbf{x} is the distance vector from the center. For $|\mathbf{v}| \ll c$ the second, magnetic part of the Lorentz force may be neglected. Under this assumption, calculate the electric dipole moment $\mathbf{p} = e\mathbf{x}$ of the electron in the direction of the external field for an oscillating field $\mathbf{E} = \mathbf{E}_0 \exp(-i\omega t)$. Assume that the dipole oscillates with the same frequency, $\mathbf{p} = \mathbf{p}_0 \exp(-i\omega t)$. Discuss the frequency dependence of the resulting polarizability.

Solution: Inserting the dipole moment into the above equation and assuming that the velocity is sufficiently small (neglecting the magnetic part of the Lorentz force), we get

$$(45.74) \quad \ddot{\mathbf{p}} + \gamma\dot{\mathbf{p}} + \omega_0^2\mathbf{p} = \frac{e^2}{m}\mathbf{E}_0 e^{-i\omega t}.$$

Further assuming that the distance vector (and the dipole moment) oscillates with the same frequency like the external electric field (driven oscillator) leads to

$$(45.75) \quad (-\omega^2 - i\omega\gamma + \omega_0^2)\mathbf{p}_0 = \frac{e^2}{m}\mathbf{E}_0$$

or

$$(45.76) \quad \mathbf{p}_0 = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \mathbf{E}_0 \equiv \alpha(\omega) \mathbf{E}_0$$

with the atomic polarizability α . The complex nature of the polarizability means that there exists a phase difference between \mathbf{E} and \mathbf{p} . This phase difference is 0 for low frequency and π in the limit of very high frequency. The real and imaginary parts of α are

$$(45.77) \quad \text{Re}\alpha = \frac{e^2}{m} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}$$

$$(45.78) \quad \text{Im}\alpha = \frac{e^2}{m} \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}.$$

This shows that there is anomalous dispersion (weaker refraction for light with higher frequency) for values of ω near the eigenfrequency ω_0 .

If there are n atoms per volume and f_k denotes the fraction of electrons per atom that oscillate with eigenfrequency ω_k , then the total polarization per volume is

$$(45.79) \quad \mathbf{p} = \frac{ne^2}{m} \sum_k \frac{f_k}{\omega_k^2 - \omega^2 - i\omega\gamma} \mathbf{E},$$

and, with $\mathbf{p} = \chi_e \mathbf{E}$, $\epsilon = 1 + 4\pi\chi_e$, we find for the electric permittivity ϵ

$$(45.80) \quad \epsilon = 1 + \frac{4\pi ne^2}{m} \sum_k \frac{f_k}{\omega_k^2 - \omega^2 - i\omega\gamma},$$

which is the formula of Drude for the permittivity. For many materials $\mu \sim 1$ and $n \sim \sqrt{\epsilon}$. For frequencies near the eigenfrequencies anomalous dispersion occurs.

45.8 Exercise 11: Fresnel's Formulae for Reflection and Refraction

Two homogeneous, isotropic transparent media (with constant, real permittivities ϵ_1, ϵ_2 and permeabilities μ_1, μ_2) are separated by a plane. Use the Maxwell equations in macroscopic matter, the resulting boundary conditions, and the plane wave ansatz to derive the laws of reflection and refraction.

Solution: The Maxwell equations in macroscopic matter, but without further microscopic charge and current distributions, are

$$(45.81) \quad \nabla \cdot \mathbf{D} = 0$$

$$(45.82) \quad \nabla \cdot \mathbf{B} = 0$$

$$(45.83) \quad \nabla \times \mathbf{E} = -\frac{1}{c} \dot{\mathbf{B}}$$

$$(45.84) \quad \nabla \times \mathbf{H} = \frac{1}{c} \dot{\mathbf{D}}$$

where

$$(45.85) \quad \mathbf{B} = \mu \mathbf{H} \quad , \quad \mathbf{D} = \epsilon \mathbf{E}.$$

Now we assume that we have a medium 1 with permittivity ϵ_1 and permeability μ_1 in the region $z < 0$, and a medium 2 with permittivity ϵ_2 and permeability μ_2 in the region $z > 0$, so the separating plane is the (x, y) plane $z = 0$. Then the following continuity conditions for the fields infinitesimally below and above the separating plane follow from the Maxwell equations:

$$(45.86) \quad \mathbf{D}_{\perp}^{(2)} = \mathbf{D}_{\perp}^{(1)}$$

$$(45.87) \quad \mathbf{B}_{\perp}^{(2)} = \mathbf{B}_{\perp}^{(1)}$$

$$(45.88) \quad \mathbf{E}_{\parallel}^{(2)} = \mathbf{E}_{\parallel}^{(1)}$$

$$(45.89) \quad \mathbf{H}_{\parallel}^{(2)} = \mathbf{H}_{\parallel}^{(1)}.$$

Here, e.g., $\mathbf{D}_{\perp}^{(2)}$ is the component perpendicular to the separating plane of the \mathbf{D} field infinitesimally above the plane (i.e., in region 2), etc.

To prove Eq. (45.86), we integrate the Maxwell equation (45.81) over the volume of a cylinder, such that the bottom and top of the cylinder are parallel discs of the same radius infinitesimally below and above the plane $z = 0$, and the cylinder mantle is perpendicular to the plane and has infinitesimal height. We get

$$(45.90) \quad \begin{aligned} 0 &= \int d^3x \nabla \cdot \mathbf{D} = \left(\int_{\text{bottom}} + \int_{\text{mantle}} + \int_{\text{top}} \right) d\mathbf{f} \cdot \mathbf{D} \\ &= \int_{\text{disc}} d\mathbf{f} (\mathbf{n}_2 + \mathbf{n}_1) \cdot \mathbf{D}, \end{aligned}$$

where we used the facts that the mantle area is infinitesimal and that the two discs have the same area. The two discs have opposite orientation, therefore $\mathbf{n}_2 = -\mathbf{n}_1$,

$$(45.91) \quad 0 = \int_{\text{disc}} d\mathbf{f} \cdot (\mathbf{D}_2 - \mathbf{D}_1),$$

and Eq. (45.86) follows (n_i are perpendicular to the plane $z = 0$). The proof of Eq. (45.87) is analogous.

To prove Eq. (45.88), we integrate Eq. (45.83) over a rectangle with corners at $(x_1, y_1, -\delta)$, (x_1, y_1, δ) , $(x_2, y_2, -\delta)$, and (x_2, y_2, δ) . Here δ is infinitesimal, whereas the distance $l = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}$ is finite. We find

$$(45.92) \quad \int_{\text{rect}} df \mathbf{n} \cdot \nabla \times \mathbf{E} = -\frac{1}{c} \int_{\text{rect}} df \mathbf{n} \cdot \nabla \times \dot{\mathbf{B}}$$

where \mathbf{n} is the unit vector perpendicular to the rectangle. Here the r.h.s. may be neglected, because the area of the rectangle is infinitesimal,

$$(45.93) \quad \lim_{\delta \rightarrow 0} \int_{\text{rect}} df \mathbf{n} \cdot \nabla \times \dot{\mathbf{B}} = 0.$$

The l.h.s. is then

$$(45.94) \quad 0 = \int_{\text{rect}} df \mathbf{n} \cdot \nabla \times \mathbf{E} = \int d\mathbf{l} \cdot (\mathbf{E}^{(2)} - \mathbf{E}^{(1)}),$$

where $d\mathbf{l}$ is the line element along the upper finite side of the rectangle, and we neglected the line integrals along the infinitesimal sides. As the orientation of the finite sides of the rectangle in the (x, y) plane is arbitrary, Eq. (45.88) follows. The proof of eq. (45.87) is analogous.

Within a given medium (i.e., with fixed ϵ and μ), the Maxwell equations may be re-written like

$$(45.95) \quad \nabla \cdot \mathbf{E} = 0$$

$$(45.96) \quad \nabla \cdot \mathbf{B} = 0$$

$$(45.97) \quad \nabla \times \mathbf{E} = -\frac{1}{c} \dot{\mathbf{B}}$$

$$(45.98) \quad \nabla \times \mathbf{B} = \frac{n}{c} \dot{\mathbf{E}}$$

where n stands for the index of refraction:

$$(45.99) \quad n \equiv \sqrt{\epsilon\mu}.$$

As for the Maxwell equations in vacuum it follows that \mathbf{E} (as well as \mathbf{B} , \mathbf{D} and \mathbf{H}) obeys the wave equation

$$(45.100) \quad \Delta \mathbf{E} - \frac{n^2}{c^2} \ddot{\mathbf{E}} = 0$$

with propagation velocity $v = c/n \leq c$. Therefore, we may use the plane wave ansatz for the electromagnetic field in a region with fixed index of refraction,

$$(45.101) \quad \mathbf{E} = \mathbf{E}_0 e^{i\mathbf{k}\mathbf{x} - i\omega t}$$

where the dispersion relation is $|\mathbf{k}| = (n/c)\omega$.

Now we want to study reflection and refraction. We assume that an incident plane wave in region 1 ($z < 0$) propagates toward the separating plane $z = 0$ (i.e., $k_z > 0$), and we want to determine a further plane wave in region 1 with $k'_z < 0$ ("reflected wave") and a plane wave in region 2 ("refracted wave") from the boundary conditions (45.86)–(45.87). Therefore we make the ansatz

$$(45.102) \quad \mathbf{E} = \mathbf{E}_0 e^{i\mathbf{k}\mathbf{x} - i\omega t}, \quad \mathbf{k} \cdot \mathbf{E}_0 = 0, \quad z < 0$$

$$(45.103) \quad \mathbf{E}' = \mathbf{E}'_0 e^{i\mathbf{k}'\mathbf{x} - i\omega' t}, \quad \mathbf{k}' \cdot \mathbf{E}'_0 = 0, \quad z > 0$$

$$(45.104) \quad \mathbf{E}'' = \mathbf{E}''_0 e^{i\mathbf{k}''\mathbf{x} - i\omega'' t}, \quad \mathbf{k}'' \cdot \mathbf{E}''_0 = 0, \quad z < 0$$

where $|\mathbf{k}| = (n_1/c)\omega$, $|\mathbf{k}'| = (n_2/c)\omega'$ and $|\mathbf{k}''| = (n_1/c)\omega''$. The other fields may be computed easily from the Maxwell equation (45.82)

$$(45.105) \quad \mathbf{B} = \frac{c}{\omega} \mathbf{k} \times \mathbf{E},$$

$$(45.106) \quad \mathbf{B}' = \frac{c}{\omega'} \mathbf{k}' \times \mathbf{E}',$$

$$(45.107) \quad \mathbf{B}'' = \frac{c}{\omega''} \mathbf{k}'' \times \mathbf{E}'',$$

and from relations (45.85).

The total solution in region 1 is now $\mathbf{E}^{(1)} = \mathbf{E} + \mathbf{E}''$, and the solution in region 2 is $\mathbf{E}^{(2)} = \mathbf{E}'$, and analogously for the other fields. Now we may use the continuity conditions (45.86)–(45.87). As they must hold everywhere along the $z = 0$ plane and for all times, the phase functions multiplying the constant vectors must be the same, i.e.,

$$(45.108) \quad \omega = \omega' = \omega'',$$

$$(45.109) \quad k_x = k'_x = k''_x, \quad k_y = k'_y = k''_y.$$

Therefore, all three wave vectors must lie in the same plane (“plane of incidence”), and we may choose without loss of generality $k_y = k'_y = k''_y = 0$ (incidence plane = (x, z) plane). The dispersion relations are now

$$(45.110) \quad |\mathbf{k}| = |\mathbf{k}''| = (n_1/c)\omega, \quad |\mathbf{k}'| = (n_2/c)\omega$$

and imply

$$(45.111) \quad k''_z = -k_z.$$

Defining the angles between the wave vectors and the z axis,

$$(45.112) \quad \cos \alpha = \frac{k_z}{|\mathbf{k}|}, \quad \cos \alpha' = \frac{k'_z}{|\mathbf{k}'|}, \quad \cos \alpha'' = -\frac{k''_z}{|\mathbf{k}''|},$$

we have therefore $\alpha = \alpha''$, i.e., the incidence angle equals the reflection angle. We further have

$$(45.113) \quad \frac{\sin \alpha}{\sin \alpha'} = \frac{k_x/|\mathbf{k}|}{k'_x/|\mathbf{k}'|} = \frac{n_2}{n_1},$$

which is the law of Snellius.

With the help of the unit vector \hat{e}_z (which is perpendicular to the plane $z = 0$) the continuity conditions (45.86)–(45.87) may be expressed like follows: for the tangential component of \mathbf{E}

$$(45.114) \quad (\mathbf{E}_0 + \mathbf{E}_0'' - \mathbf{E}'_0) \times \hat{e}_z = 0,$$

normal component of \mathbf{B}

$$(45.115) \quad (\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}_0'' - \mathbf{k}' \times \mathbf{E}'_0) \cdot \hat{e}_z = 0,$$

tangential component of $\mathbf{H} = (1/\mu)\mathbf{B}$

$$(45.116) \quad \left(\frac{1}{\mu_1} (\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}_0'') - \frac{1}{\mu_2} \mathbf{k}' \times \mathbf{E}'_0 \right) \times \hat{e}_z = 0,$$

normal component of $\mathbf{D} = \epsilon \mathbf{E}$

$$(45.117) \quad (\epsilon_1(\mathbf{E}_0 + \mathbf{E}_0'') - \epsilon_2 \mathbf{E}_0') \cdot \hat{e}_z = 0.$$

For a further investigation we specialize to two linearly independent, linearly polarized incident waves. This is sufficient, because a general incident electromagnetic plane wave may always be expressed as a linear combination of these two. The first case is a vector \mathbf{E}_0 perpendicular to the incidence plane, i.e., $\mathbf{E}_0 = E_0 \hat{e}_y$. First, we find that \mathbf{E}_0' and \mathbf{E}_0'' also have only components perpendicular to the incidence plane, $\mathbf{E}_0' = E_0' \hat{e}_y$, $\mathbf{E}_0'' = E_0'' \hat{e}_y$. For assume that $\mathbf{E}_i' = E_i' \hat{k}' \times \hat{e}_y$ (and analogous for \mathbf{E}'') are nonzero. Then we find from (45.114)

$$(45.118) \quad (E_i'' \hat{k}'' \times \hat{e}_y - E_i' \hat{k}' \times \hat{e}_y) \times \hat{e}_z = 0,$$

$$(45.119) \quad -E_i'' \cos \alpha - E_i' \cos \alpha' = 0 \quad \Rightarrow \quad E_i'' = -\frac{\cos \alpha'}{\cos \alpha} E_i'.$$

$\cos \alpha$ and $\cos \alpha'$ are greater than zero by definition, therefore the two amplitudes have different sign. On the other hand, we find from (45.115)

$$(45.120) \quad \left(\frac{E_i''}{\mu_1} \mathbf{k}'' \times (\hat{k}'' \times \hat{e}_y) - \frac{E_i'}{\mu_2} \mathbf{k}' \times (\hat{k}' \times \hat{e}_y) \right) \times \hat{e}_z = 0,$$

$$(45.121) \quad \left(-\frac{E_i''}{\mu_1} |\mathbf{k}''| + \frac{E_i'}{\mu_2} |\mathbf{k}'| \right) \hat{e}_y \times \hat{e}_z = 0,$$

$$(45.122) \quad -\frac{E_i'' n_1}{\mu_1} + \frac{E_i' n_2}{\mu_2} = 0 \quad \Rightarrow \quad E_i'' = \frac{n_2 \mu_1}{n_1 \mu_2} E_i',$$

which implies that the two amplitudes must have the same sign. This requires that both amplitudes are zero.

For the perpendicular components we find from (45.114)

$$(45.123) \quad (E_0 + E_0'' - E_0') \hat{e}_y \times \hat{e}_z = 0 \quad \Rightarrow \quad E_0 + E_0'' - E_0' = 0,$$

and from (45.116)

$$(45.124) \quad \begin{aligned} & \left(\frac{1}{\mu_1} (E_0 \mathbf{k} \times \hat{e}_y + E_0'' \mathbf{k}'' \times \hat{e}_y) - \frac{1}{\mu_2} E_0' \mathbf{k}' \times \hat{e}_y \right) \times \hat{e}_z = 0, \\ & \left(\frac{1}{\mu_1} (E_0 \mathbf{k} \cdot \hat{e}_z + E_0'' \mathbf{k}'' \cdot \hat{e}_z) - \frac{1}{\mu_2} E_0' \mathbf{k}' \cdot \hat{e}_z \right) \hat{e}_y = 0, \\ & \frac{n_1}{\mu_1} \cos \alpha (E_0 - E_0'') - \frac{n_2}{\mu_2} \cos \alpha' E_0' = 0. \end{aligned}$$

Eq. (45.115) gives the same result like Eq. (45.114), and Eq. (45.117) is trivially zero. Inserting (45.123) into (45.124) leads to

$$\begin{aligned} \sqrt{\frac{\epsilon_1}{\mu_1}} \cos \alpha (E_0 - E_0'') &= \sqrt{\frac{\epsilon_2}{\mu_2}} \cos \alpha' (E_0 + E_0''), \\ \frac{E_0''}{E_0} &= \frac{\sqrt{\epsilon_1/\mu_1} \cos \alpha - \sqrt{\epsilon_2/\mu_2} \cos \alpha'}{\sqrt{\epsilon_1/\mu_1} \cos \alpha + \sqrt{\epsilon_2/\mu_2} \cos \alpha'} = \end{aligned}$$

$$(45.125) \quad \frac{1 - \sqrt{\frac{\epsilon_2 \mu_1 \cos \alpha'}{\epsilon_1 \mu_2 \cos \alpha}}}{1 + \sqrt{\frac{\epsilon_2 \mu_1 \cos \alpha'}{\epsilon_1 \mu_2 \cos \alpha}}} = \frac{1 - \frac{\mu_1 \tan \alpha'}{\mu_2 \tan \alpha}}{1 + \frac{\mu_1 \tan \alpha'}{\mu_2 \tan \alpha}}.$$

Further,

$$(45.126) \quad \frac{E'_0}{E_0} = 1 + \frac{E''_0}{E_0} = \frac{2}{1 + \sqrt{\frac{\epsilon_2 \mu_1 \cos \alpha'}{\epsilon_1 \mu_2 \cos \alpha}}} = \frac{2}{1 + \frac{\mu_1 \tan \alpha'}{\mu_2 \tan \alpha}}.$$

For $\mu_1 = \mu_2$ (which is true for many substances) we find

$$(45.127) \quad \frac{E''_0}{E_0} = \frac{1 - \frac{\sin \alpha \cos \alpha'}{\cos \alpha \sin \alpha'}}{1 + \frac{\sin \alpha \cos \alpha'}{\cos \alpha \sin \alpha'}} = \frac{\sin(\alpha - \alpha')}{\sin(\alpha + \alpha')},$$

$$(45.128) \quad \frac{E'_0}{E_0} = \frac{2}{1 + \frac{\sin \alpha \cos \alpha'}{\cos \alpha \sin \alpha'}} = \frac{2 \cos \alpha \sin \alpha'}{\sin(\alpha + \alpha')},$$

which are the Fresnel formulae.

The second case to study is for \mathbf{E}_0 parallel to the incidence plane, $\mathbf{E}_0 = E_0 \hat{k} \times \hat{e}_y$. Again, it may be proven easily that both \mathbf{E}'_0 and \mathbf{E}''_0 must be parallel to the incidence plane, as well, in this case, $\mathbf{E}'_0 = E'_0 \hat{k}' \times \hat{e}_y$ and $\mathbf{E}''_0 = E''_0 \hat{k}'' \times \hat{e}_y$. We find from (45.114)

$$(45.129) \quad \begin{aligned} & (E_0 \hat{k} \times \hat{e}_y + E''_0 \hat{k}'' \times \hat{e}_y - E'_0 \hat{k}' \times \hat{e}_y) \times \hat{e}_z = 0, \\ & (E_0 \hat{k} \cdot \hat{e}_z + E''_0 \hat{k}'' \cdot \hat{e}_z - E'_0 \hat{k}' \cdot \hat{e}_z) \hat{e}_y = 0, \\ & (E_0 - E''_0) \cos \alpha - E'_0 \cos \alpha' = 0, \end{aligned}$$

and from (45.116)

$$(45.130) \quad \begin{aligned} & \left(\frac{1}{\mu_1} [E_0 \mathbf{k} \times (\hat{k} \times \hat{e}_y) + E''_0 \mathbf{k}'' \times (\hat{k}'' \times \hat{e}_y)] - \frac{1}{\mu_2} E'_0 \mathbf{k}' \times (\hat{k}' \times \hat{e}_y) \right) \times \hat{e}_z = 0, \\ & \left(\frac{1}{\mu_1} [E_0 |\mathbf{k}| + E''_0 |\mathbf{k}''|] - \frac{1}{\mu_2} E'_0 |\mathbf{k}'| \right) \hat{e}_y \times \hat{e}_z = 0, \\ & \frac{n_1}{\mu_1} (E_0 + E''_0) - \frac{n_2}{\mu_2} E'_0 = 0 \quad \rightarrow \quad E'_0 = \sqrt{\frac{\epsilon_1 \mu_2}{\epsilon_2 \mu_1}} (E_0 + E''_0) \end{aligned}$$

(again (45.115) and (45.117) do not contain additional information). Inserting this into the above equation leads to

$$(45.131) \quad \begin{aligned} (E_0 - E''_0) \cos \alpha &= \sqrt{\frac{\epsilon_1 \mu_2}{\epsilon_2 \mu_1}} (E_0 + E''_0) \cos \alpha', \\ \frac{E''_0}{E_0} &= \frac{\cos \alpha - \sqrt{\epsilon_1 \mu_2 / \epsilon_2 \mu_1} \cos \alpha'}{\cos \alpha + \sqrt{\epsilon_1 \mu_2 / \epsilon_2 \mu_1} \cos \alpha'} = \\ &= \frac{1 - \sqrt{\frac{\epsilon_1 \mu_2 \cos \alpha'}{\epsilon_2 \mu_1 \cos \alpha}}}{1 + \sqrt{\frac{\epsilon_1 \mu_2 \cos \alpha'}{\epsilon_2 \mu_1 \cos \alpha}}} = \frac{1 - \frac{\epsilon_1 \tan \alpha'}{\epsilon_2 \tan \alpha}}{1 + \frac{\epsilon_1 \tan \alpha'}{\epsilon_2 \tan \alpha}}. \end{aligned}$$

Further,

$$(45.132) \quad \frac{E'_0}{E_0} = \sqrt{\frac{\epsilon_1 \mu_2}{\epsilon_2 \mu_1}} \left(1 + \frac{E''_0}{E_0} \right) = \frac{2 \sqrt{\frac{\epsilon_1 \mu_2}{\epsilon_2 \mu_1}}}{1 + \frac{\epsilon_1 \tan \alpha'}{\epsilon_2 \tan \alpha}}.$$

For $\mu_1 = \mu_2$ these formulae simplify to

$$(45.133) \quad \frac{E_0''}{E_0} = \frac{1 - \frac{n_1^2 \tan \alpha}{n_2^2 \tan \alpha'}}{1 + \frac{n_1^2 \tan \alpha}{n_2^2 \tan \alpha'}} = \frac{1 - \frac{\sin \alpha' \cos \alpha'}{\sin \alpha \cos \alpha}}{1 - \frac{\sin \alpha' \cos \alpha'}{\sin \alpha \cos \alpha}} = \frac{\sin \alpha \cos \alpha - \sin \alpha' \cos \alpha'}{\sin \alpha \cos \alpha + \sin \alpha' \cos \alpha'}$$

$$(45.134) \quad \frac{E_0'}{E_0} = \frac{2 \frac{n_1}{n_2}}{1 + \frac{n_1^2 \tan \alpha}{n_2^2 \tan \alpha'}} = \frac{2 \frac{\sin \alpha'}{\sin \alpha}}{1 + \frac{\sin \alpha' \cos \alpha'}{\sin \alpha \cos \alpha}} = \frac{2 \sin \alpha' \cos \alpha}{\sin \alpha \cos \alpha + \sin \alpha' \cos \alpha'}$$

which are the Fresnel formulae for the case of electric field parallel to the incidence plane.

From these results some physical applications immediately follow. For $n_2 < n_1$ total reflection for $\alpha > \alpha_{\max}$ follows from the Snellius law, where $\sin \alpha_{\max} = n_2/n_1$.

Another interesting consequence is the possibility to generate linearly polarized light from unpolarized light by choosing the so-called Brewster angle for the incident light beam. Indeed, it follows from (45.133) that the component parallel to the incidence plane is not reflected at all if the condition

$$(45.135) \quad \sin \alpha \cos \alpha = \sin \alpha' \cos \alpha'$$

holds. The “solution” $\alpha = \alpha'$ is of course forbidden if $n_2 \neq n_1$. But there exists the possible solution $\alpha' = \pi/2 - \alpha$ which leads to $\sin \alpha' = \cos \alpha$, $\cos \alpha' = \sin \alpha$ and

$$(45.136) \quad \frac{n_2}{n_1} = \frac{\sin \alpha}{\sin \alpha'} = \frac{\sin \alpha}{\cos \alpha} = \tan \alpha.$$

Therefore, if an unpolarized light beam is incident with an angle α_B such that $\tan \alpha_B = n_2/n_1$, then the reflected beam will contain only light linearly polarized into the direction perpendicular to the incidence plane, $\mathbf{E}_0'' \sim \hat{e}_y$.

45.9 Exercise 12: Classical Zeemann Effect

Again electrons in matter are described by spherical harmonic oscillators which oscillate around a fixed center (nucleus), as in Exercise 45.7. Here we assume that the electrons are excited with their eigenfrequency by some unspecified mechanism (e.g. heat or radiation), therefore we ignore a possible damping. These harmonic oscillators are now exposed to a constant, external magnetic field, therefore the equation of motion for the position of one electron is

$$(45.137) \quad m(\ddot{\mathbf{x}}(t) + \omega_0 \mathbf{x}(t)) = \frac{e}{c} \dot{\mathbf{x}} \times \mathbf{B}.$$

Find the three eigenfrequencies and eigenmodes of this equation of motion. (Hint: use the symmetry of the problem and the known behavior of a charged particle in a constant magnetic field to guess the right ansatz.) Use the resulting dipole moments $\mathbf{p} = e\mathbf{x}$ to calculate the dipole radiation of all three modes and their polarization patterns both parallel and perpendicular to the constant, external magnetic field.

Solution: We assume without loss of generality that the constant, external magnetic field points into the positive z direction, $\mathbf{B} = B\hat{e}_3$, $B > 0$. Then we have the following equations of motion,

$$(45.138) \quad \ddot{x}_1 + \omega_0^2 x_1 = \frac{eB}{mc} \dot{x}_2,$$

$$(45.139) \quad \ddot{x}_2 + \omega_0^2 x_2 = -\frac{eB}{mc} \dot{x}_1,$$

$$(45.140) \quad \ddot{x}_3 + \omega_0^2 x_3 = 0.$$

Obviously, the third, decoupled equation is solved by $x_3 = x_0 \sin(\omega_0 t - \phi_0)$ where ϕ_0 is an irrelevant integration constant and is set to zero in the sequel. Therefore the first eigenmode is

$$(45.141) \quad \mathbf{x}_{||} = x_0 \sin(\omega_0 t) \hat{e}_3$$

with eigenfrequency ω_0 (the eigenfrequency is not altered by the magnetic field). Now we use the following facts about a charge moving in a constant magnetic field: a charge moving parallel to the magnetic field feels no force at all, whereas a charge moving perpendicular to the magnetic field is deflected such that it moves along a circle with constant angular velocity. Further, the free two-dimensional harmonic oscillator composed of Eqs. (45.138) and (45.139) has circles with constant angular velocity as possible eigenmodes, therefore we try a circle with constant angular velocity as an ansatz:

$$(45.142) \quad x_1 = x_0 \cos \omega t \quad , \quad \dot{x}_1 = -x_0 \omega \sin \omega t \quad , \quad \ddot{x}_1 = -x_0 \omega^2 \cos \omega t,$$

$$(45.143) \quad x_2 = x_0 \sin \omega t \quad , \quad \dot{x}_2 = x_0 \omega \cos \omega t \quad , \quad \ddot{x}_2 = -x_0 \omega^2 \sin \omega t.$$

We find from (45.138)

$$(45.144) \quad -x_0 \omega^2 \cos \omega t + \omega_0^2 x_0 \cos \omega t = \frac{eB}{mc} x_0 \omega \cos \omega t$$

and an analogous equation with $\cos \omega t \rightarrow \sin \omega t$ from (45.139). Therefore the ansatz is a solution provided that

$$(45.145) \quad -\omega^2 + \omega_0^2 = \frac{eB}{mc} \omega$$

$$(45.146) \quad \Rightarrow \quad \omega_{\pm} = -\frac{eB}{2mc} \pm \sqrt{\omega_0^2 + \frac{e^2 B^2}{4m^2 c^2}} \simeq -\frac{eB}{2mc} \pm \omega_0,$$

where in the last step we assumed that the eigenfrequency ω_0 is much larger in absolute value than the Larmor frequency $\omega_L = -\frac{eB}{2mc}$. The minus sign for ω_- means that this eigenmode rotates in a clockwise direction, and the minus sign of ω_L is the law of induction. Therefore, we find the following two perpendicular eigenmodes

$$(45.147) \quad \mathbf{x}_1 = x_0 \cos(\omega_1 t) \hat{e}_1 + x_0 \sin(\omega_1 t) \hat{e}_2$$

$$(45.148) \quad \mathbf{x}_2 = x_0 \cos(\omega_2 t) \hat{e}_1 - x_0 \sin(\omega_2 t) \hat{e}_2$$

where $\omega_1 = \omega_+$ and $\omega_2 = -\omega_-$.

Next, we want to calculate the dipole radiation emitted by these three eigenmodes. Here we use the formulae for Hertzian dipole radiation (see Exercise 45.6), and $\mathbf{p} = e\mathbf{x}$,

$$(45.149) \quad \mathbf{B}(t, \mathbf{x}) = \frac{e}{c^2 r^2} \ddot{\mathbf{p}}(t - \frac{r}{c}) \times \mathbf{x},$$

$$(45.150) \quad \mathbf{E}(t, \mathbf{x}) = \frac{e}{c^2 r^3} [-r^2 \ddot{\mathbf{p}}(t - \frac{r}{c}) + (\mathbf{x} \cdot \ddot{\mathbf{p}}(t - \frac{r}{c})) \mathbf{x}],$$

$$(45.151) \quad \mathbf{S} = \frac{e^2}{4\pi c^3 r^5} [r^2 \ddot{\mathbf{p}}^2 - (\mathbf{x} \cdot \ddot{\mathbf{p}})^2] \mathbf{x}.$$

For the parallel eigenmode, $\mathbf{p}_{||} = ex_0 \sin(\omega_0 t) \hat{e}_3$, we find linearly polarized radiation with polarization direction parallel to the external \mathbf{B} field (the radiation \mathbf{B} field is always perpendicular to the \hat{e}_3 direction). The linear polarization is extracted most easily from the Poynting vector

$$(45.152) \quad \mathbf{S} = \frac{e^4 \omega_0^4 x_0^2}{4\pi c^3 r^5} (r^2 - z^2) \sin^2(\omega_0 t') \mathbf{x} = \frac{e^4 \omega_0^4 x_0^2}{4\pi c^3 r^5} \sin^2 \theta \sin^2(\omega_0 t') \mathbf{n},$$

where we used spherical polar coordinates, $\mathbf{n} \equiv (\mathbf{x}/r)$, and $t' \equiv t - (r/c)$. For linear polarization at each point in space in the radiation zone the electric (and magnetic) field vector oscillates and vanishes for certain times (instead of rotating along an ellipse or a circle), therefore also the energy flow vanishes at certain times. Further, there is no radiation in the \hat{e}_3 direction.

For the eigenmodes perpendicular to the constant, external magnetic field the Poynting vector is (where $\omega = \omega_{\pm}$)

$$(45.153) \quad \mathbf{S} = \frac{e^4 \omega^4 x_0^2}{4\pi c^3 r^5} [r^2 - (x \cos(\omega t') + y \sin(\omega t'))^2] \mathbf{x} = \frac{e^4 \omega^4 x_0^2}{4\pi c^3 r^2} [1 - \sin^2 \theta \cos^2(\omega t' - \varphi)] \mathbf{n}.$$

In the (x, y) plane perpendicular to \hat{e}_3 we get (with $\mathbf{n}_{\perp} = \cos \varphi \hat{e}_1 + \sin \varphi \hat{e}_2$)

$$(45.154) \quad \mathbf{S} = \frac{e^4 \omega^4 x_0^2}{4\pi c^3 r^2} \sin^2(\omega t' - \varphi) \mathbf{n}_{\perp}$$

and, therefore, linearly polarized light. Further, the polarization direction is perpendicular to the \hat{e}_3 direction (because the magnetic radiation field is parallel to \hat{e}_3). For the radiation in \hat{e}_3 direction we get

$$(45.155) \quad \mathbf{S} = \frac{e^4 \omega^4 x_0^2}{4\pi c^3 r^2} \hat{e}_3$$

and the energy flow does not depend on time. Therefore, the radiation is circularly polarized (the radiation field vectors rotate without changing their length). Further, the $+$ mode has left circular polarization (the \mathbf{E} field rotates counter-clockwise) and the $-$ mode has right circular polarization.

45.10 Exercise 13: Diamagnetism and Paramagnetism

The relation between the magnetic induction \mathbf{B} , the magnetic field intensity \mathbf{H} and the magnetization \mathbf{M} of matter is (for isotropic matter)

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}, \quad \mathbf{B} = \mu\mathbf{H}, \quad \mathbf{M} = \chi_m\mathbf{H}, \quad \mu = 1 + 4\pi\chi_m.$$

Here a substance is called diamagnetic if $\mu < 1$. Use again the harmonic oscillator model in a constant external magnetic field (like in Exercise 45.9) to derive the diamagnetic behavior. Here the electrons moving on circular orbits have to be interpreted as currents which induce a magnetic field. Assume that the net induced magnetic field without external magnetic field is zero.

Solution: We again assume that the external magnetic field is along the \hat{e}_3 direction, $\mathbf{B} = B\hat{e}_3$. Further, we are interested only in the d.o.f. perpendicular to the external magnetic field. We know from Exercise 45.9 that there are two degrees of freedom, $+$, $-$, with frequencies

$$(45.156) \quad \omega_{\pm} = -\frac{eB}{2mc} \pm \omega_0.$$

Without external field these are just two modes rotating with the same angular velocity ω_0 , one (the $+$ mode) in the counter-clockwise direction, the other in the clockwise direction. The two modes induce magnetic moments of equal strengths but opposite orientations, therefore, macroscopically, their net contribution to the magnetization is zero. With external field, however, the additional contribution is in the clockwise (negative) direction in both cases, and a net

contribution remains. Further, the electron current and induced magnetic field depend linearly on ω , therefore we may just omit the ω_0 piece and use just $\omega_{\mathcal{L}} = -\frac{eB}{2mc}$. The magnetic moment of a current distribution is

$$(45.157) \quad \mathbf{m} = \frac{1}{2c} \int d^3x \mathbf{x} \times \mathbf{j},$$

where \mathbf{j} is the current density. In our case the electron is in a circular orbit in the (x, y) plane, therefore

$$(45.158) \quad \mathbf{j} = e\delta(\mathbf{x} - \mathbf{x}(t))\dot{\mathbf{x}}$$

(with $x_3 = 0$). With

$$(45.159) \quad \mathbf{x} \times \dot{\mathbf{x}} = (x_1\dot{x}_2 - x_2\dot{x}_1)\hat{e}_3 = x_0^2\omega_{\mathcal{L}}\hat{e}_3$$

we get

$$(45.160) \quad \mathbf{m} = \frac{1}{2c}e\omega_{\mathcal{L}}x_0^2\hat{e}_3 = -\frac{e^2Bx_0^2}{4mc^2}\hat{e}_3,$$

and, therefore, the magnetic moment is opposite to the external magnetic field. For general orbits of the electrons, x_0^2 must be replaced by the orbit average $\overline{x_1^2 + x_2^2}$; if the force law for the electron is spherically symmetric, as in our case, then $\overline{x_1^2 + x_2^2} = (2/3)\overline{\mathbf{x}^2}$. If there are several electrons per atom with different average orbit radii and if there are n atoms per volume, then the magnetization is

$$(45.161) \quad \mathbf{M} = -n\frac{e^2B}{6mc^2}\sum_j\overline{\mathbf{x}_j^2}\hat{e}_3 = -n\frac{e^2}{6mc^2}\sum_j\overline{\mathbf{x}_j^2}\mathbf{B}$$

and the susceptibility is approximately (if it is small)

$$(45.162) \quad \chi_m = -n\frac{e^2}{6mc^2}\sum_j\overline{\mathbf{x}_j^2}.$$

The assumptions here were rather general (electrons orbiting around nuclei) and, in fact, all materials are diamagnetic, but the diamagnetism may be over-compensated by other effects with positive magnetic susceptibility (like paramagnetism).

Paramagnetism is present when the following conditions hold: 1) the atoms or molecules of the material already have a fixed magnetic moment \mathbf{m}_0 (even without external magnetic field), and 2) the orientations of these magnetic moments are randomly distributed in the absence of an external magnetic field. Without external magnetic field there is, therefore, no macroscopic magnetization of the paramagnetic material. However, in an external magnetic field the magnetic moments are partially aligned along the direction of the external field, because this aligned position is energetically favorable. This alignment is thwarted by thermal fluctuations, and the average magnetic moment per molecule is

$$(45.163) \quad \overline{\mathbf{m}} = \eta\frac{|\mathbf{m}_0|^2}{kT}\mathbf{B}_{\text{ext}},$$

where η is some constant depending on the molecule type, T is the temperature and k the Boltzmann constant. Both quantum mechanical and thermodynamic considerations are needed to derive equation (45.163), which is beyond the scope of the present discussion.

45.11 Exercise 14: Stark Effect

A hydrogen atom is exposed to a constant electric field. Use perturbation theory to calculate the energy shifts for the ground state ψ_{100} of the hydrogen atom (with no energy degeneracy of the unperturbed system) and for the states ψ_{2lm} (with a four-fold energy degeneracy).

Solution: The Hamiltonian for the electron wave function in a Coulomb field and in a constant electric field is

$$(45.164) \quad H = -\frac{\hbar^2}{2\mu}\Delta - \frac{e^2}{r} - e\mathbf{E} \cdot \mathbf{x},$$

and we assume $\mathbf{E} = f\hat{e}_3$. The eigenfunctions $\psi_{nlm} = R_{nl}(r)F_l^m(\theta)e^{im\varphi}$ of the hydrogen atom are no longer eigenfunctions of the full Hamiltonian, therefore perturbation theory in the perturbing term $V = -efz = -efr\cos\theta$ is needed.

For the ground state, no degeneracy of the unperturbed system occurs, because there is only one ground state ψ_{100} with energy $E_{100} = -R$. In this case, the perturbation series for the energy is

$$(45.165) \quad E = E_k^0 + V_{kk} - \sum_{i \neq k} \frac{|V_{ki}|^2}{E_i^0 - E_k^0} + \dots$$

where E_i^0 are the eigenenergies of the unperturbed Hamiltonian $H_0 = H - V$, $H_0\psi_i = E_i^0\psi_i$, and the perturbation about the level k is calculated. Further,

$$(45.166) \quad V_{ki} = \int d^3x \overline{\psi_k(\mathbf{x})} V \psi_i(\mathbf{x})$$

is the matrix element of the perturbation V w.r.t. the eigenstates of the unperturbed system. In our case $\psi_k = \psi_{100}$,

$$(45.167) \quad V_{ki} = -ef \int_0^\infty dr r^3 \overline{R_{10}}(r) R_{nl}(r) \int d\Omega \cos\theta \overline{F_0^0}(\theta) F_l^m(\theta) e^{im\varphi},$$

which is non-zero only for $m = 0, l = 1$. Therefore, V_{kk} is zero, and there is no first order (linear in the external field) contribution. Up to second order in the perturbation, we get

$$(45.168) \quad E = -R - e^2 f^2 \text{const},$$

which is quadratic in the external field (quadratic Stark effect). For the electric dipole moment we find

$$(45.169) \quad \mathbf{p} = e\langle \mathbf{x} \rangle = -\frac{\partial}{\partial \mathbf{E}} \langle V \rangle = -\frac{\partial}{\partial \mathbf{E}} E = \text{const } e^2 \mathbf{E}.$$

The interpretation of this result is as follows: the spherically symmetric ground state ψ_{100} has no permanent electric dipole moment. But the external electric field induces a dipole moment proportional to its strength by polarizing the atom (displacing the charge center of the electron wave function relative to the nucleus).

For higher states (e.g., $n = 2$) degeneracy occurs because all states with quantum number n have energy $-(R/n^2)$. E.g. for $n = 2$ there are four states

$$(45.170) \quad \psi_1 \equiv \psi_{200}, \quad \psi_2 \equiv \psi_{210}, \quad \psi_3 \equiv \psi_{211}, \quad \psi_4 \equiv \psi_{21-1},$$

with unperturbed energy $E_2^0 = -(R/4)$. When degeneracy occurs, the energies within the degenerate subspace have to be determined exactly from the degeneracy condition

$$(45.171) \quad \det \left((E_2^0 - E)\delta_{\alpha\beta} + V_{\beta\alpha} \right) = 0,$$

where α labels the degeneracy subspace. Here

$$(45.172) \quad V_{\beta\alpha} = -ef \int_0^\infty dr r^3 \overline{R_{2l_\beta}}(r) R_{2l_\alpha}(r) \int d\Omega \cos\theta \overline{F_{l_\beta}^{m_\beta}}(\theta) F_{l_\alpha}^{m_\alpha}(\theta) e^{i(m_\alpha - m_\beta)\varphi}.$$

It follows that

$$(45.173) \quad m_\alpha = m_\beta = 0, \quad l_\beta = l_\alpha \pm 1$$

and, therefore, only V_{12} and V_{21} are non-zero. They may be calculated and one finds

$$(45.174) \quad V_{12} = V_{21} = \dots = 3efa_0,$$

where a_0 is the Bohr radius. The determinant condition is

$$(45.175) \quad \det \begin{pmatrix} E_2^0 - E & V_{12} & 0 & 0 \\ V_{12} & E_2^0 - E & 0 & 0 \\ 0 & 0 & E_2^0 - E & 0 \\ 0 & 0 & 0 & E_2^0 - E \end{pmatrix} = 0,$$

$$(45.176) \quad (E_2^0 - E)^2 [(E_2^0 - E)^2 - V_{12}^2] = 0,$$

with the solutions

$$(45.177) \quad E = E_2^0$$

(two-fold degenerate) and

$$(45.178) \quad E \equiv E_\pm = E_2^0 \pm V_{12}.$$

This expression is linear in the external field (linear Stark effect). The electric dipole moment for the two non-degenerate modes E_\pm is

$$(45.179) \quad \mathbf{p} = \frac{\partial}{\partial \mathbf{E}} E_\pm = \pm 3ea_0 \hat{e}_3.$$

The interpretation is as follows: The spherically non-symmetric wave functions give rise to a permanent electric dipole moment.

45.12 Exercise 15: Vector Model for Spin-Orbital Interaction

A phenomenological *vector model* has been used in the *old quantum theory* for an addition of the *magnetic moments* in many-electron atoms and molecules. It is not rigorous and contradicts the Pauli equation, however, sometimes it gives exact results.

Precession of angular momentum

Let us consider the orbital angular momentum \mathbf{L} .

Theorem 45.8 *In the uniform magnetic field $\mathbf{B} = (0, 0, B)$ the components $\mathbf{L}_3(t)$ and $\mathbf{s}_3(t)$ are conserved, while the vectors $(\mathbf{L}_1(t), \mathbf{L}_2(t)) \in \mathbb{R}^2$ and $(\mathbf{s}_1(t), \mathbf{s}_2(t)) \in \mathbb{R}^2$ rotate with angular velocity $\omega_{\mathcal{L}}$ and $2\omega_{\mathcal{L}}$, respectively, where $\omega_{\mathcal{L}} = -\frac{eB}{2\mu c}$ is the Larmor frequency.*

Proof For concreteness we consider the orbital momentum. The conservation of \mathbf{L}_3 is already proved. Since $\hat{\mathbf{L}}_k = -\hbar \mathbf{H}_k$, we have by (9.2),

$$(45.180) \quad [\hat{\mathbf{L}}_1, \hat{\mathbf{L}}_2] = i\hbar \hat{\mathbf{L}}_3, \quad [\hat{\mathbf{L}}_2, \hat{\mathbf{L}}_3] = i\hbar \hat{\mathbf{L}}_1, \quad [\hat{\mathbf{L}}_3, \hat{\mathbf{L}}_1] = i\hbar \hat{\mathbf{L}}_2.$$

Therefore,

$$(45.181) \quad [\mathcal{P}, \hat{\mathbf{L}}_1] = -i\omega_{\mathcal{L}}\hbar\hat{\mathbf{L}}_2, \quad [\mathcal{P}, \hat{\mathbf{L}}_2] = i\omega_{\mathcal{L}}\hbar\hat{\mathbf{L}}_1.$$

since all $\hat{\mathbf{L}}_k$ commute with all $\hat{\mathbf{s}}_j$ by Remark 19.2. Hence, analogously to the Heisenberg equation (6.25), we get

$$(45.182) \quad \begin{aligned} \dot{\mathbf{L}}_1(t) &= \langle \Psi(t), i\hbar^{-1}[\mathcal{P}\hat{\mathbf{L}}_1 - \hat{\mathbf{L}}_1\mathcal{P}]\Psi(t) \rangle \\ &= \langle \Psi(t), \omega_{\mathcal{L}}\hat{\mathbf{L}}_2\Psi(t) \rangle = \omega_{\mathcal{L}}\mathbf{L}_2(t). \end{aligned}$$

Similarly, we have $\dot{\mathbf{L}}_2(t) = -\omega_{\mathcal{L}}\mathbf{L}_1(t)$, hence

$$(45.183) \quad \frac{d}{dt}(\mathbf{L}_1(t) + i\mathbf{L}_2(t)) = -i\omega_{\mathcal{L}}(\mathbf{L}_1(t) + i\mathbf{L}_2(t)). \quad \blacksquare$$

The theorem implies the *precession* of the vectors $\mathbf{L}(t)$ and $\mathbf{s}(t)$ with the angular velocities $\omega_{\mathcal{L}}$ and $2\omega_{\mathcal{L}}$, respectively, around the magnetic field \mathbf{B} .

Remark 45.9 A similar precession can be proved for a classical system of electrons rotating as rigid bodies in the uniform magnetic field [53].

Vector model

Let us apply the idea of the precession to the addition of the orbital and spin magnetic moments. This addition explains the Einstein-de Haas and anomalous Zeemann experiments.

The model explains the magnetization in the Einstein-de Haas experiment by the classical mechanism of the reorientation in a magnetic field of a total magnetic moment \mathbf{m} which exists even in the absence of the magnetic field. \mathbf{m} is the sum of the orbital and spin magnetic moments $\mathbf{m}_o := \frac{e}{2\mu c}\mathbf{L}$

and $\mathbf{m}_s := \frac{e}{\mu c}\mathbf{s}$:

$$(45.184) \quad \mathbf{m} = \mathbf{m}_o + \mathbf{m}_s = \frac{e}{2\mu c}\mathbf{L} + \frac{e}{\mu c}\mathbf{s}.$$

The moments are defined by the orbital and spin angular momenta \mathbf{L} and \mathbf{s} , respectively, with the corresponding distinct gyromagnetic ratios $\frac{e}{2\mu c}$ and $\frac{e}{\mu c}$.

In the absence of an external magnetic field, the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{s}$ is conserved, while \mathbf{L} and \mathbf{s} are generally not conserved. For example, the spin angular momentum \mathbf{s} precesses in the magnetic field generated by the orbital angular momentum. Similarly, the orbital angular momentum \mathbf{L} precesses in the magnetic field generated by the spin angular momentum. Therefore, the lengths of the vectors \mathbf{L} and \mathbf{s} are conserved. Hence, the conservation of the sum $\mathbf{J} = \mathbf{L} + \mathbf{s}$ implies that the vectors \mathbf{L} and \mathbf{s} rotate around \mathbf{J} . Then the total magnetic moment $\mathbf{m} = \frac{e}{2\mu c}\mathbf{L} + \frac{e}{\mu c}\mathbf{s}$ also rotates around \mathbf{J} . Since the angular velocity of the rotation is very high, we have to take into account only the effective value of the magnetic moment, \mathbf{m}_{eff} , which is the projection of the total magnetic moment onto \mathbf{J} . Let us calculate this projection.

The angle α between the vectors \mathbf{J} and \mathbf{L} is conserved as well as the angle β between the vectors \mathbf{J} and \mathbf{s} , and

$$(45.185) \quad \cos \alpha = \frac{|\mathbf{J}|^2 + |\mathbf{L}|^2 - |\mathbf{s}|^2}{2|\mathbf{J}||\mathbf{L}|}, \quad \cos \beta = \frac{|\mathbf{J}|^2 + |\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}||\mathbf{s}|}.$$

Then the projection equals

$$(45.186) \quad \mathbf{m}_{\text{eff}} = \frac{e}{2\mu c}|\mathbf{L}| \cos \alpha + \frac{e}{\mu c}|\mathbf{s}| \cos \beta.$$

A final, highly nontrivial approximation is as follows: we redefine the lengths of the vectors \mathbf{J} , \mathbf{L} and \mathbf{s} as

$$(45.187) \quad |\mathbf{J}|^2 := \langle \Psi, (\hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2) \Psi \rangle, \quad |\mathbf{L}|^2 := \langle \Psi, (\hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2) \Psi \rangle, \quad |\mathbf{s}|^2 := \langle \Psi, (\hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2) \Psi \rangle.$$

The operators $\hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$, $\hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2$, and $\hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2 = 3/4$ commute. Hence, the quantum stationary states can be classified by the eigenvalues of the operators which are equal to $J(J+1)$, $L(L+1)$, and $3/4$, where $J, L = 0, 1, 2, \dots$. For the states we have $|\mathbf{J}|^2 = J(J+1)$, $|\mathbf{L}|^2 = L(L+1)$, and $|\mathbf{s}|^2 = 3/4$. Substituting this into (45.186), we obtain *the Landé formula* (20.8) for the effective gyromagnetic ratio

$$(45.188) \quad g_{\text{eff}} := \frac{\mathbf{m}_{\text{eff}}}{|\mathbf{J}| \frac{e}{2\mu c}} = \frac{|\mathbf{J}|^2 + |\mathbf{L}|^2 - |\mathbf{s}|^2}{2|\mathbf{J}|^2} + 2 \frac{|\mathbf{J}|^2 + |\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}|^2} \\ = \frac{3}{2} + \frac{|\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}|^2} = \frac{3}{2} + \frac{3/4 - L(L+1)}{2J(J+1)}.$$

The formula is confirmed experimentally by the Einstein-de Haas and anomalous Zeemann effects.

Bibliography

- [1] M.Abraham, Theorie der Elektrizität, Bd.2: Elektromagnetische Theorie der Strahlung, Teubner, Leipzig, 1905.
- [2] V.Arnold, Mathematical Methods of Classical Mechanics, Springer, Berlin, 1978.
- [3] A.V.Babin, M.I.Vishik, Attractors of Evolutionary Equations, North-Holland, Amsterdam, 1992.
- [4] R.Becker, Elektronentheorie, Teubner, Leipzig, 1933.
- [5] A.Bensoussan, C.Iliine, A.Komech, Breathers for a relativistic nonlinear wave equation, *Arch. Rat. Mech. Anal.* **165** (2002), 317-345.
- [6] H.Berestycki, P.L.Lions, *Arch. Rat. Mech. and Anal.* **82** (1983), no.4, 313-375.
- [7] F.A.Berezin, M.A.Shubin, The Schrödinger equation, Kluwer Academic Publishers, Dordrecht, 1991.
- [8] H.Bethe, Intermediate Quantum Mechanics, Benjamin, NY, 1964.
- [9] H.Bethe, E.Salpeter, Quantum Mechanics of One- and Two-Electron Atoms, Berlin, Springer, 1957.
- [10] J.D.Bjorken, S.D.Drell, Relativistic Quantum Mechanics, McGraw-Hill, NY, 1964; Relativistic Quantum Fields, McGraw-Hill, NY, 1965.
- [11] N.Bohr, *Phil. Mag.* **26** (1913) 1; 476; 857.
- [12] M.Born, Atomic Physics, Blackie & Son, London-Glasgow, 1951.
- [13] N.Bournaveas, Local existence for the Maxwell-Dirac equations in three space dimensions, *Commun. Partial Differ. Equations* **21** (1996), no.5-6, 693-720.
- [14] L. de Broglie, The Current Interpretation of Wave Mechanics, Elsevier, NY, 1964.
- [15] V.S.Buslaev, G.S.Perelman, On the stability of solitary waves for nonlinear Schrödinger equations, *Amer. Math. Soc. Trans. (2)* **164** (1995), 75-98.
- [16] V.S.Buslaev, C.Sulem, On asymptotic stability of solitary waves for nonlinear Schrödinger equations, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **20** (2003), no.3, 419-475.
- [17] E.U.Condon, E.U., G.H.Shortley, The Theory of Atomic Spectra, Cambridge University Press, Cambridge, 1963.
- [18] S.Cuccagna, Stabilization of solutions to nonlinear Schrödinger equations, *Comm. Pure Appl. Math.* **54** (2001), no.9, 1110-1145.

- [19] S.Cuccagna, On asymptotic stability in 3D of kinks for the ϕ^4 model, preprint 2004.
- [20] C.Davisson, L.Germer, *Nature* **119** (1927), 558.
- [21] M.Defranceschi, C.Le Bris, *Mathematical Models and Methods for ab Initio Quantum Chemistry*, Lecture Notes in Chemistry. 74, Springer, Berlin, 2000.
- [22] T.Dudnikova, A.Komech, E.A.Kopylova, Yu.M.Suhov, On convergence to equilibrium distribution, I. Klein-Gordon equation with mixing, *Comm. Math. Phys.* **225** (2002), no.1, 1-32.
- [23] T.Dudnikova, A.Komech, N.Mauser, On the convergence to a statistical equilibrium for the Dirac equation, *Russian Journal of Math. Phys.* **10** (2003), no.4, 399-410.
- [24] T.Dudnikova, A.Komech, N.Mauser, On two-temperature problem for harmonic crystals, *Journal of Statistical Physics* **114** (2004), no.3/4, 1035-1083.
- [25] T.Dudnikova, A.Komech, N.E.Ratanov, Yu.M.Suhov, On convergence to equilibrium distribution, II. Wave equation with mixing, *Journal of Statistical Physics* **108** (2002), no.4, 1219-1253.
- [26] T.Dudnikova, A.Komech, H.Spohn, On a two-temperature problem for wave equation with mixing, *Markov Processes and Related Fields* **8** (2002), no.1, 43-80.
- [27] T.Dudnikova, A.Komech, H.Spohn, On convergence to statistical equilibrium for harmonic crystals, *Journal of Mathematical Physics* **44** (2003), no.6, 2596-2620.
- [28] T.Dudnikova, A.Komech, H.Spohn, Energy-momentum relation for solitary waves of relativistic wave equation, *Russian Journal Math. Phys.* **9** (2002), no.2, 153-160.
- [29] D.M.Eidus, The principle of limit amplitude, *Russ. Math. Surv.* **24** No.3, 97-167 (1969).
- [30] A.Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (Concerning the generation and transformation of light as seen from a heuristic point of view), *Annalen der Physik*, March 18, 1905.
- [31] M.Esteban, V.Georgiev, E.Sere, Stationary solutions of the Maxwell-Dirac and the Klein-Gordon-Dirac equations, *Calc. Var. Partial Differ. Equ.* **4** (1996), no.3, 265-281.
- [32] M. Fedoriuk, *Asymptotic Methods for Partial Differential Equations*, Encyclopaedia of Mathematical Sciences Vol. 34, Springer, Berlin, 1999.
- [33] R.Feynman, *Quantum Electrodynamics*, Addison-Wesley, Reading, Massachusetts, 1998.
- [34] Ya.Frenkel, *Zs. Phys.* **37** (1926), 43.
- [35] J.Fröhlich, S.Gustafson, B.L.G.Jonsson, I.M.Sigal, Solitary wave dynamics in an external potential, arXiv:math-ph/0309053v1.
- [36] J.Fröhlich, T.P.Tsai, H.T.Yau, On the point-particle (Newtonian) limit of the nonlinear Hartree equation, *Comm. Math. Physics* **225** (2002), no.2, 223-274.
- [37] I.Gelfand, S.Fomin, *Calculus of Variations*, Dover Publications, Mineola, NY, 2000.
- [38] I.M.Gelfand, R.A.Minlos, Z.Ya.Shapiro, *Representations of the Rotation and Lorentz Groups and their Applications*, Pergamon Press, Oxford, 1963.

- [39] M.Gell-Mann, *Phys. Rev.* **125** (1962), 1067.
- [40] M.Grillakis, J.Shatah, W.A.Strauss, Stability theory of solitary waves in the presence of symmetry, I; II. *J. Func. Anal.* **74** (1987), no.1, 160-197; **94** (1990), no.2, 308-348.
- [41] S.J.Gustafson, I.M.Sigal, *Mathematical Concepts of Quantum Mechanics*, Springer, Berlin, 2003.
- [42] Y.Guo, K.Nakamitsu, W.Strauss, Global finite-energy solutions of the Maxwell-Schrödinger system, *Comm. Math. Phys.* **170** (1995), no. 1, 181–196.
- [43] J.Hale, *Asymptotic Behavior of Dissipative Systems*, AMS, Providence, 1988.
- [44] K.Hannabus, *An Introduction to Quantum Theory*, Clarendon Press, Oxford, 1997.
- [45] W.Heisenberg, Der derzeitige Stand der nichtlinearen Spinortheorie der Elementarteilchen, *Acta Phys. Austriaca* **14** (1961), 328-339.
- [46] W.Heitler, *The Quantum Theory of Radiation*, Oxford University Press, Oxford, 1954.
- [47] D.Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics, 840, Springer, Berlin, 1981.
- [48] P.D.Hislop, I.M.Sigal, *Introduction to Spectral Theory. With applications to Schrödinger operators*, Springer, NY, 1996.
- [49] V.Imaikin, A.Komech, P.Markowich, Scattering of solitons of the Klein-Gordon equation coupled to a classical particle, *Journal of Mathematical Physics* **44** (2003), no.3, 1202-1217.
- [50] V.Imaikin, A.Komech, H.Spohn, Soliton-like asymptotics and scattering for a particle coupled to Maxwell field, *Russian Journal of Mathematical Physics* **9** (2002), no.4, 428-436.
- [51] V.Imaikin, A.Komech, H.Spohn, Scattering theory for a particle coupled to a scalar field, *Journal of Discrete and Continuous Dynamical Systems* **10** (2003), no.1&2, 387-396.
- [52] C.Itzykson, J.B.Zuber, *Quantum Field Theory*, McGraw-Hill, NY, 1980.
- [53] Jackson, *Classical Electrodynamics*, 3rd ed., John Wiley & Sons, New York, 1999.
- [54] V.Jaksic, C.-A.Pillet, Ergodic properties of classical dissipative systems. I. *Acta Math.* **181** (1998), no.2, 245-282.
- [55] M.Jammer, *The Conceptual Development of Quantum Mechanics*, McGraw-Hill, NY, 1966.
- [56] P.Joly, A.Komech, O.Vacus, On transitions to stationary states in a Maxwell-Landau-Lifschitz-Gilbert system, *SIAM J. Math. Anal.* **31** (1999), no.2, 346-374.
- [57] T.Kato, A.Jensen, Spectral properties of Schrödinger operators and time-decay of the wave functions, *Duke Math. J.* **46** (1979), 583-611.
- [58] W.Kauffmann, *Ann. Physik* **61** (1897), 544.
- [59] A.Komech, On stabilization of string-nonlinear oscillator interaction, *J. Math. Anal. Appl.* **196** (1995), 384-409.
- [60] A.Komech, On transitions to stationary states in Hamiltonian nonlinear wave equations, *Phys. Letters A* **241** (1998), 311-322.

- [61] A.Komech, On transitions to stationary states in one-dimensional nonlinear wave equations, *Arch. Rat. Mech. Anal.* **149** (1999), no.3, 213-228.
- [62] A.Komech, Attractors of nonlinear Hamiltonian one-dimensional wave equations, *Russ. Math. Surv.* **55** (2000), no.1, 43-92.
- [63] A.Komech, On attractor of a singular nonlinear $U(1)$ -invariant Klein-Gordon equation, p. 599-611 in: Proceedings of the 3rd ISAAC Congress, Freie Universität Berlin, Berlin, 2003.
- [64] A.Komech, On Global Attractors of Hamilton Nonlinear Wave Equations, preprint 2005.
- [65] A.Komech, M.Kunze, H.Spohn, Effective Dynamics for a mechanical particle coupled to a wave field, *Comm. Math. Phys.* **203** (1999), 1-19.
- [66] A.Komech, H.Spohn, Soliton-like asymptotics for a classical particle interacting with a scalar wave field, *Nonlinear Analysis* **33** (1998), no.1, 13-24.
- [67] A.Komech, H.Spohn, Long-time asymptotics for the coupled Maxwell-Lorentz equations, *Comm. Partial Diff. Eqs.* **25** (2000), no.3/4, 558-585.
- [68] A.Komech, H.Spohn, M.Kunze, Long-time asymptotics for a classical particle interacting with a scalar wave field, *Comm. Partial Diff. Eqs.* **22** (1997), no.1/2, 307-335.
- [69] A.Komech, N.Mauser, A.Vinnichenko, On attraction to solitons in relativistic nonlinear wave equations, to appear in *Russian J. Math. Phys.*
- [70] A.Landé, *Zs. f. Phys.* **5** (1921), 5.
- [71] V.G.Levich, Yu.A.Vdovin, V.A.Mamlin, Course of Theoretical Physics, v. II, Nauka, Moscow, 1971. [Russian]
- [72] E.H.Lieb, The Stability of Matter: from Atoms to Stars, Springer, Berlin, 2001.
- [73] H.A.Lorentz, The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat. Lectures from a course held at Columbia University, New York, NY, USA, March and April 1906, Repr. of the 2nd ed., Sceaux: ditions Jacques Gabay, 1992.
- [74] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press Princeton, 1955.
- [75] R.Newton, Quantum Physics, Springer, NY, 2002.
- [76] E.Noether, Gesammelte Abhandlungen. Collected papers, Springer, Berlin, 1983.
- [77] S.P.Novikov, S.V.Manakov, L.P.Pitaevskii, V.E.Zakharov, Theory of Solitons: The Inverse Scattering Method, Consultants Bureau [Plenum], New York, 1984.
- [78] J.Perrin, New experiments on the cathode rays, *Nature*, **53** (1896), 298-299.
- [79] C.-A.Pillet, C.E.Wayne, Invariant manifolds for a class of dispersive, Hamiltonian, partial differential equations, *J. Differ. Equations* **141** (1997), No.2, 310-326.
- [80] R.Pyke, A.Soffer, M.I.Weinstein, Asymptotic stability of the kink soliton of the φ^4 nonlinear equation, preprint, 2003.

- [81] M.Reed, B.Simon, *Methods of Modern Mathematical Physics*, Academic Press, NY, I (1980), II (1975), III (1979), IV (1978).
- [82] J.J.Sakurai, *Advanced Quantum Mechanics*, Addison-Wesley, Reading, Massachusetts, 1967.
- [83] G.Scharf, *Finite Quantum Electrodynamics. The Causal Approach*, Springer, Berlin, 1995.
- [84] L.Schiff, *Quantum Mechanics*, McGraw-Hill, NY, 1955.
- [85] E.Schrödinger, Quantisierung als Eigenwertproblem, *Ann. d. Phys.* I, II **79** (1926) 361, 489; III **80** (1926) 437; IV. **81** (1926) 109.
- [86] M.A.Shubin, *Pseudodifferential Operators and Spectral Theory*, Springer-Verlag, Berlin, 2001.
- [87] A.Soffer, M.I.Weinstein, Multichannel nonlinear scattering for nonintegrable equations, *Comm. Math. Phys.* **133** (1990), 119-146.
- [88] A.Soffer, M.I.Weinstein, Multichannel nonlinear scattering for nonintegrable equations. II. The case of anisotropic potentials and data. *J. Differential Equations* **98** (1992) no. 2 376.
- [89] A.Sommerfeld, *Atombau und Spektrallinien*, Vol. I and II, Friedr. Vieweg & Sohn, Braunschweig, 1951.
- [90] A.Sommerfeld, G.Schur, *Ann. d. Phys.* **4** (1930), 409.
- [91] H.Spohn, *Dynamics of Charged Particles and Their Radiation Field*, Cambridge University Press, Cambridge, 2004.
- [92] W.A.Strauss, *Nonlinear Invariant Wave Equations*, Lecture Notes in Phys.73 (1978), Springer, Berlin, 197-249.
- [93] C.Sulem, P.L.Sulem, *The Nonlinear Schrödinger Equation. Self-Focusing and Wave Collapse*, Springer, NY, 1999.
- [94] A.Szabo, N.S.Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover, 1996.
- [95] R.Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Applied Mathematical Sciences, Springer, Berlin, 1988.
- [96] B.Thaller, *The Dirac Equation*, Springer, Berlin, 1991.
- [97] L.H.Thomas, *Nature* **117** (1926), 514.
- [98] J.J.Thomson, *Phil. Mag.* **44** (1897), 298.
- [99] B.R.Vainberg, *Asymptotic Methods in Equations of Mathematical Physics*, Gordon and Breach, New York, 1989.
- [100] G.Wentzel, *ZS. f. Phys.* **43** (1927), 1, 779.
- [101] H.Weyl, *The Theory of Groups and Quantum Mechanics*, Dover, NY, 1949.
- [102] E.P.Wigner, *Theory of Group*, Academic Press, NY, 1959.
- [103] E.Zeidler, *Applied Functional Analysis. Main Principles and Their Applications*, Applied Mathematical Sciences Vol.109, Springer, Berlin, 1995.