Lecture Notes Physik

Theoretical Astrophysics

An Introduction

MATTHIAS BARTELMANN

HEIDELBERG UNIVERSITY PUBLISHING

THEORETICAL ASTROPHYSICS
AN INTRODUCTION

Lecture Notes Physik

THEORETICAL ASTROPHYSICS

AN INTRODUCTION

Matthias Bartelmann

Institut für Theoretische Physik
Universität Heidelberg

HEIDELBERG
UNIVERSITY PUBLISHING

About the Author

Matthias Bartelmann https://orcid.org/0000-0001-6951-3582
is professor for theoretical astrophysics at Heidelberg University. He mostly addresses cosmological questions, concerning in particular the formation and evolution of cosmic structures.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

The electronic open access version of this work is permanently available on Heidelberg University Publishing's website: https://heiup.uni-heidelberg.de
urn: urn:nbn:de:bsz:16-heiup-book-822-9
doi: https://doi.org/10.17885/heiup.822

ISSN 2566-4816 (PDF)
ISSN 2512-4455 (Print)

ISBN 978-3-96822-029-1 (PDF)
ISBN 978-3-96822-028-4 (Softcover)

Contents

Preface

This book is not in any sense complete or exhaustive, and it is not meant to be. Its subject, theoretical astrophysics, is vast and cannot possibly be comprehensively covered in a single volume.

This book has a rather different purpose. It is intended as a textbook for students who have a reasonably complete knowledge of the material usually taught in the introductory courses on theoretical physics: classical mechanics, electrodynamics, quantum mechanics, and thermodynamics. Building upon this assumed foundation, this book adds material typically not covered by the introductory lectures, but required for research work in theoretical astrophysics. It may also be useful as a resource for researchers. Arguably the most important extensions are radiation processes, hydrodynamics, plasma physics and magnetohydrodynamics, and stellar dynamics.

This book provides introductions to these four areas. It is structured into four main chapters and an initial chapter summarising some essential theoretical concepts which the following chapters build upon.

The chapter on radiation processes begins with the Larmor equation from electrodynamics and derives Thomson scattering and a general approach to calculating spectra from it, which is then applied to synchrotron radiation and bremsstrahlung. Up to this point, electromagnetic radiation is described as a classical wave that does not exchange momentum with the charges it originates from or interacts with. The backreaction of radiation on the radiating charge is discussed then before Compton scattering is introduced, and with it the photon picture of electromagnetic radiation. The internal structure of radiating systems such as atoms follows, leading to the calculation of cross sections for the interaction of quantum-mechanical systems with radiation and of the shapes of spectral lines. Finally, radiation is described as an ensemble of photons. Specific intensity, emissivity and opacity, the Planck spectrum and radiation transport are introduced there.

The chapter on hydrodynamics begins with a derivation of the ideal hydrodynamical equations from elementary kinetic theory. It is emphasised that these equations express the (local) conservation of the energy-momentum tensor. This opens the way into relativistic hydrodynamics as well as towards various extensions, such as viscous hydrodynamics and magneto-hydrodynamics. The assumption of an infinitely small mean free path from ideal hydrodynamics is then relaxed, leading to viscous hydrodynamics. Inviscid and viscous flows are considered under certain simplifying conditions. The formation of shocks and the Sedov solution follow before the discussion of several fluid instabilities concludes the chapter. The discussion in this chapter emphasises the root of hydrodynamics in the conservation equation for the energy-momentum tensor, the common origin of non-ideal hydrodynamical effects in particle transport, the importance of integrated statements such as Kelvin's theorem, the Bernoulli equation and the Rankine-Hugoniot conditions, and the general approach to linear perturbation or stability analysis.

The chapter on plasma physics begins with the introduction of the plasma parameters and proceeds to the propagation of electromagnetic waves through a plasma. Dispersion relations are derived generally for transverse and longitudinal waves, touching the phenomenon of Landau damping, and specified for thermal plasmas. The equations of magneto-hydrodynamics are introduced next, emphasising their common ground with hydrodynamics in the vanishing divergence of an energy-momentum tensor. The generation of magnetic fields is briefly discussed, followed by ambipolar diffusion as an example for non-ideal coupling between the plasma charges and the fluid particles. The propagation of electromagnetic waves through cold, magnetised plasmas is studied next, and the chapter concludes with a linear stability analysis, revealing the variety of hydromagnetic and Alfvén modes.

The chapter on stellar dynamics begins with deriving Jeans' equations in parallel to the hydrodynamical equations, emphasising the importance of anisotropic pressure. Stability criteria for stellar-dynamical systems are then derived, leading to the Jeans and Toomre criteria. Finally, the phenomenon of dynamical friction is introduced and discussed, ending with Chandrasekhar's formula for the friction force.

Preparing for this selection of subjects, the initial chapter briefly summarises special relativity and relativistic electrodynamics as well as elementary kinetic theory to lay the foundation for the discussions in the following main chapters.

In all chapters, the attempt was made to trace these four areas of theoretical astrophysics back to their origins in fundamental concepts of theoretical physics. Rather than discussing many examples and trying to cover as many astronomical and astrophysical phenomena as possible, the goal of this book is to reveal the roots of the common approaches in theoretical astrophysics, the choices and assumptions made and the methodical similarities appearing throughout. The book does not aim at explaining the richness of astrophysical phenomena, but at enabling the reader to understand and apply the rich toolbox of theoretical astrophysics by her- or himself. In this spirit, the notorious phrases "one can show" or "as can be shown" do not appear in this book. Every subject discussed is derived from first principles, which is considerably more important to the author than completeness.

This book grew from a one-semester course in theoretical astrophysics developed and regularly taught at the University of Heidelberg. The course comprised four hours of lecture and a two-hour tutorial per week. The amount of material collected here is probably at the upper end of what can be covered in a single term of 15 weeks. If it needs to be pruned, the general idea of the course should not be given up: to reveal the foundations of theoretical astrophysics including its important general assumptions, and to identify the common methodical approaches.

By far the most, if not all of the material summarised and compiled in this book is not new. Its intention lies in the foundation and the arrangement of matters, which may help seeing them from a common and unifying perspective. It is not at all possible to give full reference to the original derivations and presentations, not to mention any specific research results. This is therefore not even attempted. Rather, we give a list of more specialised textbooks and refer to them for further reading on individual subjects.

Acknowledgments

Since this book emerged from a course in theoretical astrophysics, many students were exposed to it. I am most grateful to the very many comments and suggestions and the constructive criticism I received, which helped to clarify and sharpen the arguments, to prune derivations of unnecessary loops and details, and to remove errors. In particular, I want to name Santiago Casas, Carsten Littek, Martina Schwind, Elena Sellentin and Benjamin Wallisch, who read and commented on large parts or all of the first edition of this book before it was published.

The perpetual discussions with the members of our cosmology group, first at the Institut für Theoretische Astrophysik, now at the Institut für Theoretische Physik of Heidelberg University were extremely helpful and clarifying. I wish to thank you all, guys, for your great help and support! In particular, I should mention Christian Angrick, Matteo Maturi, Philipp Merkel, and Christophe Pixius for designing many problems, working out solutions, unfailingly criticising weaknesses and improving the line of reasoning.

The idea for a course like this dates back to discussions with my friend and colleague Achim Weiß, now almost two decades ago. I profited greatly and over many years from his unconventional and stimulating view on our common field of research. Many colleagues inspired and impressed me with their broad and comprehensive view of physics, their clarity of thinking, and their attitude both towards research and teaching. I wish to expressly name Jürgen Ehlers, Wolfgang Hillebrandt, Jens Niemeyer, Manfred Salmhofer, Björn Schäfer, Peter Schneider, Norbert Straumann, Christof Wetterich and Simon White, who shaped and influenced this book in many ways, mostly without knowing it.

Large parts of the first edition of this book were written at personally difficult times. For the support and encouragement I received during this time, I am deeply grateful to a fine group of friends and colleagues: Thank you all for kindness and attention and for precious, lovely hours!

Chapter 1

Theoretical Foundations

1.1 Units

1.1.1 Lengths, masses, times, and temperatures

We use Gaussian centimetre-gram-second (cgs) units throughout. Lengths are measured in cm, masses in grams and time in seconds. The derived units of force, energy and power are listed in Table 1.1. Temperatures are unvariedly measured in Kelvin (K).

Table 1.1 The units of force, energy and power are listed here in the cgs system together with their relations to SI units.

	quantity	cgs unit	alternatives	
force	mass · acceleration	$\frac{\mathrm{g\,cm}}{\mathrm{s}^2}$	dyn	10^{-5} N
energy	mass · velocity^2	$\frac{\mathrm{g\,cm}^2}{\mathrm{s}^2}$	erg	10^{-7} J
power	energy / time	$\frac{\mathrm{erg}}{\mathrm{s}}$		10^{-7} W

The main reason for using these rather than SI units is they allow electromagnetic relations to be expressed in a much easier way, as we shall now discuss.

1.1.2 Charges and electromagnetic fields

The unit of charge is chosen such that the Coulomb force between two charges q separated by the distance r is

$$F_{\mathrm{Coulomb}} = \frac{q^2}{r^2} \;. \tag{1.1}$$

With this choice, the dielectric constant of the vacuum, ε_0, becomes dimensionless and unity. Electric and magnetic fields are defined to have the same unit. This is most sensible in view of the fact that they are both related, and can be converted into each other, by Lorentz transforms. Their unit is chosen such that the force caused by an electric field E on a charge q is

$$F_{\text{electric}} = qE \;. \tag{1.2}$$

?

Confirm the cgs units of charge and electric or magnetic fields listed in Tab. 1.2.

This implies that charge, electric and magnetic fields must have the units given in Table 1.2. The squared electric or magnetic field strengths then have the dimension of an energy density.

Table 1.2 This table lists the units of charge, electric and magnetic field in the Gaussian cgs system, their physical dimensions, and alternative units.

quantity		cgs unit	alternative
charge	force$^{1/2}\cdot$ length	$\frac{\text{g}^{1/2}\text{cm}^{3/2}}{\text{s}}$	esu
electric or magnetic field	force / charge	$\frac{\text{g}^{1/2}}{\text{cm}^{1/2}\,\text{s}}$	Gauss

By definition, the units of charge in the SI and the Gaussian cgs systems are related by

$$1\,\text{Coulomb} = 2.9979\cdot 10^{9}\,\text{esu}\;. \tag{1.3}$$

Electrostatic potential differences, or electrostatic potential energy changes per unit charge, are measured in Volts in SI units. Consequently, we must have

$$1\,\text{Volt} = 1\,\frac{\text{Joule}}{\text{Coulomb}} = \frac{10^7\,\text{erg}}{2.9979\cdot 10^9\,\text{esu}} = \frac{1}{299.79}\,\frac{\text{g}^{1/2}\,\text{cm}^{1/2}}{\text{s}}\;. \tag{1.4}$$

The energy gained by a unit charge moving through an electrostatic potential difference of 1 Volt, defined as the electron-Volt, must then be

?

Use the Boltzmann constant k_B to convert 1 eV to an equivalent temperature.

$$1\,\text{eV} = 1.6022\cdot 10^{-12}\,\text{erg}\;. \tag{1.5}$$

1.1.3 Natural constants

The most frequently used natural constants in cgs units are tabulated in Table 1.3.

Caution Note that the light speed is exact by definition of the metre. ◄

In addition, some units used in astronomy and astrophysics are listed in Table 1.4.

1.1.4 Conventions and notation

For the Minkowski metric, we use the signature

$$\eta = \text{diag}\,(-1,+1,+1,+1)\;. \tag{1.6}$$

We adopt the convention

$$\tilde{f}(k) = \mathcal{F}\,[f]\,(k) = \int \frac{\text{d}^d k}{(2\pi)^d}\, f(x)\,\text{e}^{-\text{i}k\cdot x} \tag{1.7}$$

Table 1.3 The most frequently used natural constants are tabulated here with their common symbols and their values in cgs units. The values are taken from the Particle Data Group (http://pdg.lbl.gov/, last accessed on Nov. 22, 2020).

quantity	symbol	value in cgs units
light speed	c	$2.9979 \cdot 10^{10}$
elementary charge	e	$4.8032 \cdot 10^{-10}$
electron mass	m_e	$9.1094 \cdot 10^{-28}$
proton mass	m_p	$1.6726 \cdot 10^{-24}$
Boltzmann's constant	k_B	$1.3806 \cdot 10^{-16}$
Newton's constant	G	$6.6743 \cdot 10^{-8}$
Planck's constant	$\hbar$	$1.0546 \cdot 10^{-27}$

Table 1.4 Some units common in astronomy and astrophysics are listed here.

unit	symbol	type	value in cgs units
Solar radius	$R_\odot$	length	$6.9634 \cdot 10^{10}$
astronomical unit	AU	length	$1.4960 \cdot 10^{13}$
light year	ly	length	$9.4607 \cdot 10^{17}$
parsec	pc	length	$3.0857 \cdot 10^{18}$
Earth mass	$M_\oplus$	mass	$5.9724 \cdot 10^{27}$
Jupiter mass	$M_♃$	mass	$1.8990 \cdot 10^{30}$
Solar mass	$M_\odot$	mass	$1.9884 \cdot 10^{33}$
tropical year	y	time	$3.1557 \cdot 10^{7}$
sidereal year	y	time	$3.1558 \cdot 10^{7}$
Solar luminosity	$L_\odot$	energy/time	$3.8460 \cdot 10^{33}$
Jansky	Jy	specific intensity	10^{-23}

for the Fourier transform in d dimensions, and

$$f(x) = \mathcal{F}^{-1}(x) = \int \mathrm{d}^d x\, \tilde{f}(k)\, \mathrm{e}^{\mathrm{i}k\cdot x} \tag{1.8}$$

for its inverse. We use the short-hand notation

$$\int_x := \int \mathrm{d}^d x \quad \text{and} \quad \int_k := \int \frac{\mathrm{d}^d k}{(2\pi)^d} \tag{1.9}$$

for the integrals over coordinates $x \in \mathbb{R}^d$ and wave vectors $k \in \mathbb{R}^d$.

1.2 Lorentz Invariance

This section summarises the concepts of special relativity and their consequences for the structure of space-time and for the dynamics of a particle. Its most important results are the relativistic time dilation (1.36) and the Lorentz contraction (1.40), the addition theorem for velocities (1.42) and the transformation of angles (1.45), the combination of energy and momentum into the momentum four vector (1.63) and the relativistic relations (1.66) and (1.67) between energy, momentum and velocity.

Perhaps it is helpful to begin with the statement that classical physics aims to quantify the behaviour of physical entities in space with time. Point mechanics, for example, studies the trajectories of particles with negligible extension. A trajectory can be quantified by a vector-valued function $\vec{x}(t)$ which assigns a spatial vector $\vec{x}$ to any instant t from a finite or infinite time interval. Field theory describes forces as the effect of fields, which are functions of space and time obeying their own dynamics. Immediately, we are led to the question how we want to identify points in space and instants in time in a quantifiable manner.

This is achieved by a reference frame or a coordinate system. In Newtonian physics, space and time were both assumed to be absolute. A rigid reference frame was assumed to exist which identified each point in space by a triple $\vec{x}$ of real-valued, spatial coordinates, and by a real number t for the time. Having formulated the laws of physics in this absolute frame, the immediate further question arises as to how other frames of reference, or coordinate systems, could be chosen such that those laws would remain valid without changing their form. The answer of Newtonian physics was that the laws of physics are the same in all so-called inertial frames. In slightly different words, the laws of physics were claimed to be invariant under all transformations leading from one inertial frame to another.

A clarifying remark should be in order here before we move on. Notice the perhaps trivial point that not the physical *quantities* are generally assumed to be unchanged under transformations from one inertial frame to another, but the *form of the physical laws* relating them. For example, Newton's second axiom, force is mass times acceleration, is expected to hold in all inertial frames, irrespective of the specific values of the acceleration and the force. In another inertial frame, the values of force and acceleration may and generally will be different, but the statement of the law, force equals mass times acceleration, is expected to remain valid. Valid physical laws are expected to be invariant in this sense. If, in addition, physical quantities can be identified that remain invariant under transformations from one inertial frame to another, such conserved quantities play an important role in the analysis of specific physical systems under consideration. It is thus of central importance for any part of theoretical physics to clearly state which type of transformation should lead from one inertial frame to another.

In a more mathematical language, transformations between inertial frames form groups. Admissible physical laws are those which are invariant under the operation of those groups. The identification of the invariance group of a physical theory is perhaps the most fundamental step in its foundation.

1.2.1 The Special Lorentz Transform

In Newtonian mechanics, inertial frames are related by Galilei transformations. If one inertial frame is given, any Galilei transform turns it into another one. The Galilei transforms form a ten-parameter group of transformations. They contain shifts of the origin in space and time (four parameters), translations with constant velocity (three parameters), and rotations in space (further three parameters, e.g. the Euler angles). Consequences of the Galilei invariance of Newtonian mechanics are the existence of an absolute time and the Galilean addition theorem for velocities.

However, the Galilei invariance of Newtonian mechanics leads to contradictions with experience. The decay of muons sets a prominent example. Myons are leptons comparable to the electron, but with a mass of 105.6 MeV instead of 0.511 MeV. They decay according to

$$\mu \to e^- + \bar{\nu}_e + \nu_\mu \tag{1.10}$$

into electrons and (anti-) neutrinos with a half-life of $\tau_\mu = 1.5 \cdot 10^{-6}\,\mathrm{s}$. Experiments show, however, that the lifetime increases if the muon moves in the laboratory frame with velocities near the speed of light. The electron emitted in the decay has almost light speed, but never exceeds it even if the muon had already moved with almost the speed of light. Clearly, the muon seems to live longer in the laboratory rest frame than in its own rest frame, and the Galilean theorem for adding velocities does not longer apply.

Einstein's theory of Special Relativity replaced the Galilei invariance of Newtonian mechanics by the Lorentz invariance of relativistic physics. Special Relativity grew from the problem that the speed of light c appears as an absolute velocity in Maxwell's vacuum equations of electrodynamics. Einstein radically solved this problem by elevating the postulate to a principle that the speed of light c is a universal constant, independent of the state of motion of the light source relative to the observer. Interestingly, the concepts of absolute space and time underlying Newtonian physics were thus replaced by the concept of an absolute, observer-independent maximal velocity.

Consider now two inertial frames, S and S', moving relative to each other at an arbitrary, constant speed (Figure 1.1). Imaginge a flash of light going off. By the principle of the constant light speed, the wave front of the flash must

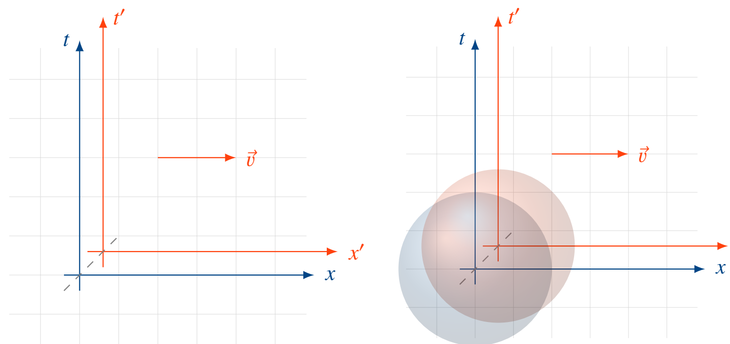

Figure 1.1 *Left*: Two inertial frames are shown moving with constant velocity $\vec{v}$ relative to each other. They are synchronised such that their origins coincide at $t = 0 = t'$. *Right*: A light signal emerging from a source at the common origin of both frames, illustrated by the coloured spheres, propagates in the same way in both frames, despite the relative motion of the two frames.

propagate in the same way in both frames irrespective of their relative velocity and therefore obey the condition

$$d\vec{x}^2 - c^2 dt^2 = d\vec{x}'^2 - c^2 dt'^2 \ . \tag{1.11}$$

For definiteness and without loss of generality, we now rotate the coordinate frames S and S' such that they move with respect to each other along their common $\hat{e}_z$ axis, and further set the origin of time such that both frames coincide at $t = 0 = t'$. Requiring further that the transformation between S and S' be linear leads directly to the *special Lorentz transform*

$$x'^3 = \gamma\left(x^3 + \beta ct\right) \ , \quad ct' = \gamma\left(ct + \beta x^3\right) \ , \tag{1.12}$$

where $\beta = v/c$ is the relative velocity in units of the light speed, and the *Lorentz factor*

$$\gamma := \left(1 - \beta^2\right)^{-1/2} \tag{1.13}$$

appears. In the limit of low velocities, $\beta \ll 1$, the Lorentz factor is $\gamma \approx 1 + \beta^2/2$ to second order in β, or $\gamma \approx 1$ to first order. Note that we write the vector indices in (1.12) as superscripts. This may appear arbitrary here, but has a deeper mathematical sense that will shortly be explained.

As (1.12) shows, the time t and the spatial coordinates x^i cannot be uniquely or invariantly separated under special Lorentz transforms. They lose their independent identity and become coupled to each other, depending on the relative motion of the frames in which they are measured. Instead of the rigid Newtonian, Euclidean space-time with its uniquely defined, absolute time axis, we thus need to adopt a four-dimensional space-time with a different structure. We introduce $ct := x^0$ as a further coordinate and combine the coordinate quadruples to *four-vectors* $x = (x^\mu) = (x^0, x^1, x^2, x^3)^\top$. This four-dimensional space with a structure to be clarified below is called Minkowski space $\mathbb{M} = \mathbb{R}^{3+1}$.

The Lorentz transform connects any two inertial frames in the four-dimensional Minkowski space. General Lorentz transforms are composed of special Lorentz transforms in all spatial directions, the so-called Lorentz boosts, plus the orthogonal three-dimensional spatial rotations. Poincaré transforms are general Lorentz transforms combined with arbitrary translations in space and time. Like the Galilei transformations, Poincaré transformations have ten parameters: the three Euler angles for the orthogonal three-dimensional rotations, the four translations, and one velocity for the Lorentz boosts in all three independent spatial directions. In relativistic mechanics, the Poincaré transformations replace the Galilei transformations of Newtonian mechanics.

1.2.2 Minkowski Space

Since Lorentz transforms leave the expression $-(x^0)^2 + \vec{x}^2$ invariant by construction, we define the Minkowskian scalar product between two four-vectors as

$$\langle x, y\rangle = -x^0 y^0 + \vec{x}\cdot\vec{y} = \eta(x, y) \,, \tag{1.14}$$

where $\vec{x}\cdot\vec{y}$ is the ordinary scalar product between two vectors in Euclidean space. The product $\langle\cdot,\cdot\rangle$ is a pseudo-scalar product because it is not positive semi-definite. Based on this scalar product, the Lorentz group as the invariance group of relativistic physics, abbreviated by $O(3, 1)$, can now formally be defined as the set of all linear transforms represented by real-valued, square, 4×4 matrices $\mathcal{M}(4, \mathbb{R})$ that leave the scalar product (1.14) unchanged,

$$O(3, 1) = \{\Lambda \in \mathcal{M}(4, \mathbb{R}) : \langle \Lambda x, \Lambda y\rangle = \langle x, y\rangle \;\forall\; x, y \in \mathbb{M}\} \,. \tag{1.15}$$

?

Show that the condition

$$\Lambda^\top \eta \Lambda = \eta$$

is equivalent to $\langle \Lambda x, \Lambda y\rangle = \langle x, y\rangle$.

This clearly repeats as a mathematical statement that Lorentz transforms are defined as those linear transforms leaving the speed of light invariant.

The object η introduced in (1.14) satisfies the definition of a second-rank tensor, as it is a bilinear map of two vectors from Minkowski space $\mathbb{M}$ into the real numbers,

$$\eta : \mathbb{M}\times\mathbb{M} \to \mathbb{R} \,, \quad (x, y) \mapsto \eta(x, y) = \langle x, y\rangle \,. \tag{1.16}$$

This tensor is the metric tensor of Minkowski space, or the Minkowski metric. Generally, a metric is a second-rank, symmetric tensor which is non-degenerate. This means that if $\langle x, y\rangle = 0$ for all $x \in \mathbb{M}$, then $y = 0$. Once a Cartesian coordinate basis is introduced for Minkowski space, the metric can be represented by the diagonal matrix

$$\left(\eta_{\mu\nu}\right) = \mathrm{diag}(-1, 1, 1, 1) \,, \tag{1.17}$$

which allows us to write the scalar product (1.14) as

$$\langle x, y\rangle = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu y^\nu \,. \tag{1.18}$$

The subscripted indices introduced here are again not arbitrarily set and will be further illustrated below. By means of the metric, the linear map x^* defined by

$$x^* : \mathbb{M} \to \mathbb{R} \,, \quad y \mapsto x^*(y) = \eta(x, y) = \langle x, y\rangle \tag{1.19}$$

can be introduced on Minkowski space. It maps vectors into the real numbers as shown. The set of all such linear maps forms the dual vector space $\mathbb{M}^*$ to Minkowski space $\mathbb{M}$.

While vector components are identified by upper indices, dual-vector components are written with lower indices. Then, according to

$$\langle x, y\rangle = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu y^\nu = \sum_{\nu=0}^{4} \left(\sum_{\mu=0}^{4} \eta_{\mu\nu} x^\mu \right) y^\nu \ , \tag{1.20}$$

the dual vector x^* of a four-vector x has the components

$$x_\nu = \sum_{\mu=0}^{4} \eta_{\mu\nu} x^\mu = \left(-x^0, x^1, x^2, x^3\right) \ . \tag{1.21}$$

In Euclidean space, the distinction between vectors and dual vectors is irrelevant because its metric can be represented by the unit matrix. In Minkowski space, it becomes vitally important because of the minus sign of the time-time (or 0-0) component in the metric.

We now introduce Einstein's sum convention in the following form. If an index appears twice in a product and at different levels (i.e. one sub- and one superscripted), a sum over the repeated index is implied. Thus, for example,

$$x_\mu y^\mu = \sum_{\mu=0}^{3} x_\mu y^\mu \ . \tag{1.22}$$

This notation simplifies the previous expressions considerably. Written in components, the scalar product between two vectors x and y simply becomes

$$\langle x, y\rangle = x_\mu y^\mu \ . \tag{1.23}$$

The notation of four-vectors and their dual vectors is made consistent by writing the inverse Minkowski metric η^{-1} with superscripted indices, since then

$$x^\mu = \eta^{\mu\alpha} x_\alpha = \eta^{\mu\alpha} \eta_{\alpha\nu} x^\nu = \delta^\mu_\nu x^\nu \ . \tag{1.24}$$

Thus, we must have

$$\eta^{\mu\alpha} \eta_{\alpha\nu} = \delta^\mu_\nu \ , \tag{1.25}$$

from which we conclude that the matrix representations of the Minkowski metric as well as of its inverse can be brought into the diagonal form

$$(\eta^{\mu\nu}) = \left(\eta_{\mu\nu}\right) = \mathrm{diag}(-1, 1, 1, 1) \ . \tag{1.26}$$

In the notation developed so far, the special Lorentz transform (1.12) can be written as

$$x'^\mu = \Lambda^\mu{}_\nu\, x^\nu \quad \text{with} \quad (\Lambda^\mu{}_\nu) = \begin{pmatrix} \gamma & 0 & 0 & \gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\beta & 0 & 0 & \gamma \end{pmatrix} \ . \tag{1.27}$$

Since the Lorentz transform is constructed to leave the Minkowski scalar product invariant, recall (1.15), we must have

$$\eta_{\alpha\beta} x^\alpha x^\beta = \langle x, x\rangle = \langle x', x'\rangle = \eta_{\mu\nu} \Lambda^\mu{}_\alpha\, x^\alpha \Lambda^\nu{}_\beta\, x^\beta = \left(\eta_{\mu\nu} \Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\right) x^\alpha x^\beta \ , \tag{1.28}$$

showing that the Lorentz transform also leaves the Minkowski metric invariant,

$$\eta_{\alpha\beta} = \eta_{\mu\nu}\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta} \; . \tag{1.29}$$

This relation replaces the perhaps more familiar orthonormality relation in Euclidean space. There, orthonormal transformations R need to satisfy the condition

$$(R\vec{x}) \cdot (R\vec{y}) = \vec{x} \cdot \vec{y} \; , \tag{1.30}$$

which implies the condition $R^{\top} = R^{-1}$ on matrix representations of R.

?

Compare (1.29) with the condition

$$\Lambda^{\top}\eta\Lambda = \eta \; .$$

The Minkowskian orthonormality relation (1.29) implies that dual-vector components must transform under Lorentz transformations as

$$x'_{\mu} = \Lambda_{\mu}{}^{\nu}x_{\nu} \; , \tag{1.31}$$

which differs from the transformation (1.27) of vector components. Quantities transforming like vector or dual-vector components under Lorentz transforms are called Lorentz contravariant or covariant, respectively. Quantities unchanged by Lorentz transforms are Lorentz invariant. Vectors are consequently sometimes addressed as contravariant vectors, dual vectors as covariant vectors, which is a terminology which we avoid here because it hides the more fundamental mathematical distinction between vectors and dual vectors (which is also decisively important elsewhere, e.g. in quantum mechanics).

Since the coordinate time becomes largely arbitrary in Special Relativity as it loses any invariant meaning, it needs to be replaced by an invariant measure of time. The only Lorentz-invariant quantity that can be defined to characterise the separation between two space-time points x^{μ} and $x^{\mu} + \mathrm{d}x^{\mu}$ is the so-called line element of the Minkowski metric (1.14),

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} \; . \tag{1.32}$$

This line element is interpreted as the so-called proper time $\mathrm{d}\tau$ by the identification

$$\mathrm{d}s^2 = -c^2\mathrm{d}\tau^2 \; . \tag{1.33}$$

This definition is meaningful since the proper time equals the time measured by an observer in his or her own rest frame. In that frame, an observer arbitrarily placed at the spatial coordinate origin has the Minkowski coordinates $(x^0, 0, 0, 0)^{\top}$. Two subsequent events experienced by that observer at instants of coordinate time separated by $\mathrm{d}x^0$ in the rest frame have the invariant distance

$$c^2\mathrm{d}\tau^2 = \left(\mathrm{d}x^0\right)^2 = c^2\mathrm{d}t^2 \; , \tag{1.34}$$

which shows that the proper time agrees with the coordinate time in any observer's rest frame.

1.2.3 Some Properties of the Minkowski World

We briefly summarise some essential conclusions from the Lorentz covariance of the Minkowski world (see also Figure 1.2). First, let two events happen in the unprimed system S at the same location $\vec{x} = 0$, but with a time difference

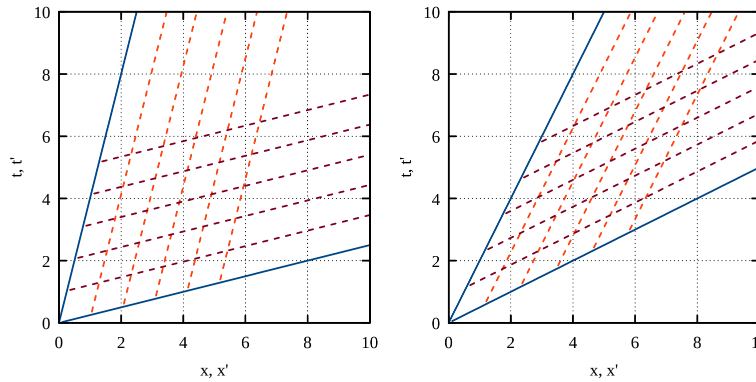

Figure 1.2 Lines of constant t' (dark red) and x' (light red) are shown in the unprimed system S for $\beta = 0.25$ (*left*) and $\beta = 0.5$ (*right*). The lines are inclined with an angle $\arctan(\beta)$ relative to the unprimed axes.

δt or $\delta x^0 = c\delta t$. These events have the four-vectors $x_1 = (0, 0, 0, 0)^\top$ and $x_2 = (\delta x^0, 0, 0, 0)^\top$. By the special Lorentz transform (1.27), they are transformed into the events

$$x'_1 = (0, 0, 0, 0)^\top \quad , \quad x'_2 = \left(\gamma\delta x^0, 0, 0, \beta\gamma\delta x^0\right)^\top \; . \tag{1.35}$$

Thus, in the primed system S', they are separated by the larger time interval

$$\delta x'^0 = \gamma\delta x^0 \quad \text{or} \quad \delta t' = c^{-1}\delta x'^0 = \gamma\delta t \; . \tag{1.36}$$

This is the relativistic time dilation: *Moving clocks go slow*.

Next, we consider a unit rule oriented in the direction of the relative motion of the two frames and resting in the unprimed system S. Its end points, measured at an arbitrary time $ct = x^0$ in S, are marked by the four-vectors $x_1 = (x_1^0, 0, 0, 0)$ and $x_2 = (x_2^0, 0, 0, 1)$. Now, an observer in S' measures its end points. It is important that he does so at one fixed instant of his coordinate time, which we arbitrarily and without loss of generality set to be $x'^0 = 0$. By (1.27), this requires

$$0 = x'^0 = \gamma x^0 + \beta\gamma x^3 \quad \text{or} \quad x^0 = -\beta x^3 \; . \tag{1.37}$$

For the two end points of our unit rule, this simultaneity condition implies that

$$x_1^0 = 0 \quad \text{and} \quad x_2^0 = -\beta \tag{1.38}$$

since $x_1^3 = 0$ and $x_2^3 = 1$ by construction. The unit rule's end points $x_{1,2}$ appear at

$$x'_1 = \left(\gamma x_1^0, 0, 0, \beta\gamma x_1^0\right)^\top \quad , \quad x'_2 = \left(\gamma x_2^0 + \beta\gamma, 0, 0, \beta\gamma x_2^0 + \gamma\right)^\top \tag{1.39}$$

in the primed observer's rest frame S'. Inserting (1.38) here gives

$$x_1'^3 = 0 \quad \text{and} \quad x_2'^3 = \left(1 - \beta^2\right)\gamma = \gamma^{-1} \; . \tag{1.40}$$

Thus, in the primed system S', the unit rule turns out to have the length $x_2'^3 - x_1'^3 = \gamma^{-1}$, which is smaller than its unit length in the rest frame. This is the relativistic length contraction: *Moving rods are shorter*.

Let us now consider a particle moving with velocity $\vec{u} = (u_x, u_y, u_z)^\top$ in the unprimed system. Its four vector in S, $x = (x^0, u_x t, u_y t, u_z t)^\top = x^0(1, u_x/c, u_y/c, u_z/c)^\top$, is transformed into

$$x' = x_0 \left(\gamma + \beta\gamma \frac{u_z}{c}, \frac{u_x}{c}, \frac{u_y}{c}, \beta\gamma + \gamma \frac{u_z}{c} \right)^\top . \tag{1.41}$$

The velocity components of the particle in the primed system S' are then found to be

$$u'_{x,y} = c\frac{x'^{1,2}}{x'^0} = \frac{u_{x,y}}{\gamma(1 + \beta u_{x,y}/c)} , \quad u'_z = c\frac{x'^3}{x'^0} = \frac{v + u_z}{1 + v u_z/c^2} . \tag{1.42}$$

The last equation is the relativistic law for the addition of velocities. While the velocity components perpendicular to the relative motion of the two frames S and S' are reduced by the Lorentz factor γ, the velocity component parallel to the motion adds to the relative velocity of the two frames in such a way that the sum of the two velocities u_z and v never exceeds c.

Let the particle now fly with the speed of light into a direction enclosing the angle θ with the $\hat{e}_z$ axis along which the two frames move relative to each other. For simplicity, but without loss of generality, we further rotate both coordinate frames about their common $\hat{e}_z$ axis such that the particle moves in the x-z coordinate plane. Then,

$$u_x = c \sin\theta , \quad u_y = 0 , \quad u_z = c \cos\theta , \tag{1.43}$$

in the unprimed system, and

$$u'_x = c \sin\theta , \quad u'_y = 0 , \quad u'_z = \frac{v + c\cos\theta}{1 + \beta\cos\theta} \tag{1.44}$$

in the primed system. Since the absolute velocity must also remain $|\vec{u}'| = c$ in the primed frame, the direction of motion in S' is

$$\cos\theta' = \frac{u'_z}{c} = \frac{\beta + \cos\theta}{1 + \beta\cos\theta} . \tag{1.45}$$

This is the relativistic aberration of light: Light rays propagating perpendicularly to $\hat{e}_z$ in S enclose an angle $\theta' = \arccos\beta$ with the $\hat{e}'_z$ axis in S'. For non-relativistic velocities, $\beta \ll 1$ and $\cos\theta' \approx \beta + \cos\theta$ to first order in β.

?

Confirm the non-relativistic limit of the relation (1.45).

Consequently, the solid-angle element spanned by light rays also changes due to the relative motion of S' relative to S. As the velocity components perpendicular to the direction of motion are unchanged, so is the azimuthal angle, $\phi' = \phi$ and $\mathrm{d}\phi' = \mathrm{d}\phi$. From the aberration formula (1.45), we have

$$\mathrm{d}\cos\theta' = \frac{\mathrm{d}\cos\theta}{1 + \beta\cos\theta} - \frac{(\beta + \cos\theta)\beta \mathrm{d}\cos\theta}{(1 + \beta\cos\theta)^2} = \frac{\mathrm{d}\cos\theta}{\gamma^2(1 + \beta\cos\theta)^2} , \tag{1.46}$$

which implies that the solid-angle element spanned by a light bundle transforms as

$$\mathrm{d}\Omega' = \mathrm{d}\phi' \mathrm{d}\cos\theta' = \frac{\mathrm{d}\phi \mathrm{d}\cos\theta}{\gamma^2(1 + \beta\cos\theta)^2} = \frac{\mathrm{d}\Omega}{\gamma^2(1 + \beta\cos\theta)^2} . \tag{1.47}$$

This is relativistic beaming: Isotropic radiation in the unprimed system S attains a highly anisotropic angular distribution in S', pointing strongly into the forward direction (Figures 1.3 and 1.4).

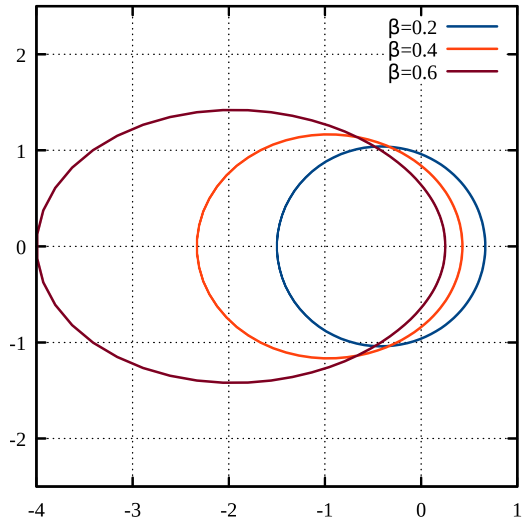

Figure 1.3 Illustration of the relativistic deformation of the solid-angle element $\mathrm{d}\Omega'/\mathrm{d}\Omega$ for the three different velocities $\beta = 0.2, 0.4, 0.6$ as indicated. The curves illustrate how isotropic radiation emitted by a point source resting in the unprimed system S would appear focussed into the direction of motion in the primed system S'.

1.2.4 Relativistic Dynamics

Since the coordinate time has no invariant meaning any more in relativity, the definition of velocity must be changed. The four-velocity is introduced as the derivative of a position four-vector with respect to the invariant proper time τ,

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \; . \tag{1.48}$$

By definition of the proper time in (1.32),

$$\begin{aligned} \mathrm{d}\tau &= c^{-1}\sqrt{-\mathrm{d}s^2} = c^{-1}\sqrt{-\mathrm{d}x^\mu \mathrm{d}x_\mu} = c^{-1}\sqrt{c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^2} = \mathrm{d}t\sqrt{1-\beta^2} \\ &= \gamma^{-1}\mathrm{d}t \; . \end{aligned} \tag{1.49}$$

Accordingly, the components of the four-velocity are

$$u^\mu = \gamma\,(c, \vec{v})^\top = c\gamma\left(1, \vec{\beta}\right)^\top \; , \tag{1.50}$$

hence its (Minkowski) square is

$$u^2 = \langle u, u\rangle = u^\mu u_\mu = -c^2\gamma^2\left(1-\beta^2\right) = -c^2 \; , \tag{1.51}$$

which is obviously and by construction invariant. Since $\mathrm{d}\tau$ is also invariant, u^μ transforms like the four-vector x^μ under Lorentz transformations, and is thus also a four-vector.

Similarly, the four-momentum of a particle with mass m is defined as

$$p^\mu = mu^\mu = \gamma mc\left(1, \vec{\beta}\right) \; . \tag{1.52}$$

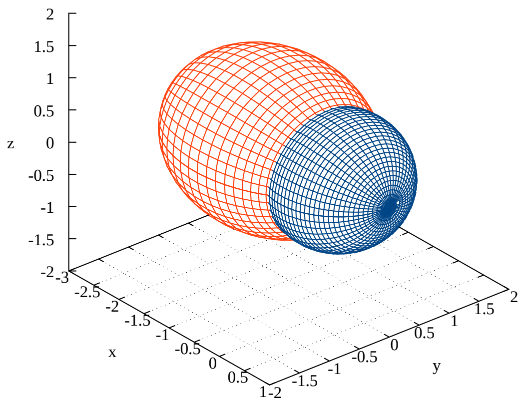

Figure 1.4 The relativistic deformation of the solid angle is shown here in a pseudo-three-dimensional representation. The blue sphere around a source at rest illustrates isotropy. When the source is moving, the sphere surrounding it in its rest frame appears strongly distorted into its forward direction.

Up to second order in β, the zero (time) component of the four-momentum is

$$p^0 = \gamma mc \approx mc\left(1 + \frac{\beta^2}{2}\right) = c^{-1}\left(mc^2 + \frac{m}{2}v^2\right) . \tag{1.53}$$

Here, the non-relativistic kinetic energy $mv^2/2$ appears together with the rest energy mc^2.

In analogy to classical mechanics, we now search for the action S of a free, relativistic particle, i.e. a particle moving relativistically in absence of external forces. The action must be Lorentz invariant since it must not depend on the arbitrary state of motion of any observer. Therefore, it must only depend on Lorentz scalars characterising a free particle. For a free particle, the only such scalar is the proper time τ, scaled with a constant α to be determined later,

$$S = \alpha \int_a^b \mathrm{d}\tau , \tag{1.54}$$

where a and b mark the fixed four-dimensional start and end points of the particle's trajectory. The action must have the dimension [energy]·[time]. Since τ has the dimension [time], the constant α must be a constant energy, which we shall determine later.

Writing the action as a function of the coordinate time t, we find

$$S = \alpha \int_{t_a}^{t_b} \mathrm{d}t \sqrt{1 - \beta^2} , \tag{1.55}$$

from which we can identify the Lagrange function

$$L(\vec{x}, \vec{v}, t) = \alpha \sqrt{1 - \beta^2} \tag{1.56}$$

for the free relativistic particle. For non-relativistic motion, $\beta \ll 1$, this must reproduce the Lagrange function of a free particle in Newtonian mechanics,

$$\alpha \sqrt{1 - \beta^2} \approx \alpha \left(1 - \frac{\beta^2}{2}\right) = \alpha - \frac{\alpha v^2}{c^2} . \tag{1.57}$$

Ignoring the irrelevant constant α left as a first term on the right-hand side, the second term shows that $\alpha = -mc^2$ must be chosen to satisfy the limit of non-relativistic mechanics. Accordingly, the action of the free relativistic particle is

$$S = -mc^2 \int_a^b \mathrm{d}\tau \,, \tag{1.58}$$

and its Lagrange function is

$$L = -mc^2 \sqrt{1-\beta^2} \,. \tag{1.59}$$

The Euler-Lagrange equation requires

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \vec{v}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma mc^2 \frac{\vec{\beta}}{c}\right) = mc\frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma\vec{\beta}\right) = 0 \,, \tag{1.60}$$

which implies $\dot{\vec{\beta}} = 0 = \dot{\vec{v}}$: The free particle moves on a straight line, as expected.

The momentum conjugate to the three-dimensional position vector $\vec{x}$ is

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1-\beta^2}} = \gamma m\vec{v} \,. \tag{1.61}$$

The particle's Hamilton function follows from the Legendre transform

$$H = \vec{v}\cdot\vec{p} - L = \gamma m v^2 + mc^2\sqrt{1-\beta^2} = \gamma mc^2\left(\beta^2 + \frac{1}{\gamma^2}\right) = \gamma mc^2 \,. \tag{1.62}$$

This is to be interpreted as the energy E of the particle. Taking the results (1.61) and (1.62) together and comparing them with the momentum four-vector shows that we can write the latter in the form

$$p^\mu = (E/c, \vec{p})^\top \,. \tag{1.63}$$

This identifies the momentum four-vector with the energy-momentum vector of a relativistic particle. Its Minkowski square is

$$\langle p, p\rangle = p_\mu p^\mu = -\frac{E^2}{c^2} + \vec{p}^2 \,, \tag{1.64}$$

while the equivalent definition $p^\mu = mu^\mu$ implies

$$\langle p, p\rangle = m^2\langle u, u\rangle = -m^2c^2 \,. \tag{1.65}$$

Together, (1.64) and (1.65) form the relativistic energy-momentum relation

$$E^2 = c^2\vec{p}^2 + m^2c^4 \,. \tag{1.66}$$

Combining (1.61) and (1.62) finally gives the very useful relation

$$\vec{p} = \frac{E}{c^2}\vec{v} = \frac{E}{c}\vec{\beta} \,. \tag{1.67}$$

Let us conclude this section with a remark on energy, momentum and their conservation in relativity. Energy and momentum are conserved if the Lagrange- or Hamilton functions of a system are invariant under translations in time and space, respectively. In relativity, time and space lose their independent existence. Time intervals and spatial distances can at least partially be transformed into each other, depending on the observer's state of motion relative to the system considered. Therefore, separate energy-momentum conservation cannot retain an invariant meaning in relativistic mechanics, and must be combined to the joint energy-momentum conservation.

Problems

1. Recall the mathematical definitions of a group, a field, a vector space, a scalar product, a dual vector space, and a tensor.
2. Write down the transformations of time $t \to t'$ and position $\vec{x} \to \vec{x}'$ under Galilei transformations.
3. Which of the following quantities are Lorentz invariant?
$$\vec{x}^2 \,, \quad x_\mu x^\mu \,, \quad x^\mu x^\nu \,, \quad \eta_{\mu\nu} \,, \quad \mathrm{d}s^2 \,, \quad \left(\mathrm{d}x^0\right)^2 \,, \quad \gamma \,, \quad \mathrm{d}\tau^2 \tag{1.68}$$
4. Compute the following expressions:
$$\partial_\alpha x^\mu \,, \quad \partial_\alpha x_\mu \,, \quad \partial_\alpha \langle x, x\rangle = \partial_\alpha \left(x^\mu x_\mu\right) \,. \tag{1.69}$$
5. Light rays are described by their wave vector $k^\mu = (\omega/c, \vec{k})$, where $\vec{k}$ is the three-dimensional wave vector pointing into the propagation direction of the light ray and satisfying the vacuum dispersion relation $\omega = ck$ with the frequency ω.
   (a) Compute the (Lorentz-invariant) scalar product of the wave vector k^μ and an arbitrary four-velocity u^μ. Explain why the frequency measured by an observer moving with four-velocity u^μ is
$$\omega_{\mathrm{obs}} = -\langle u, k\rangle = -u_\mu k^\mu \,. \tag{1.70}$$
   (b) Comparing two observers, one at rest and one moving with respect to the first with velocity $\vec{v}$, derive the relativistic Doppler relation
$$\frac{\omega'}{\omega} = \frac{1 - \vec{n}\cdot\vec{\beta}}{\sqrt{1-\beta^2}} \,, \tag{1.71}$$
   where $\vec{\beta} = \vec{v}/c$ and $\vec{n} = \vec{k}/k$.
   (c) The four-momentum of a particle is $p^\mu = mu^\mu$, where the four-velocity
$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tag{1.72}$$
   is the derivative of the coordinates x^μ with respect to the proper time τ. Starting from the relativistic Hamilton function
$$H = \frac{1}{2m} p_\mu p^\mu \,, \tag{1.73}$$
   of a free particle, derive the equations of motion and show that its Lagrange function is
$$L = \frac{m}{2} u_\mu u^\mu \,. \tag{1.74}$$
6. Beginning with the defining condition
$$\Lambda^\top \eta \Lambda = \eta \quad \text{with} \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.75}$$
   for the Lorentz transform in two dimensions,

(a) argue why an angle ψ must exist such that

$$\Lambda(\psi) = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} . \tag{1.76}$$

(b) Define $\beta = \tanh\psi$ and show that $\cosh\psi = \gamma$ and $\sinh\psi = \beta\gamma$.

(c) Show that $\Lambda(\psi_1)\Lambda(\psi_2) = \Lambda(\psi_1 + \psi_2)$. Use this result to derive the relativistic law for adding velocities.

1.3 Electromagnetism

This section summarises the foundations and some important results of classical electrodynamics. The theory is motivated as the only Lorentz invariant, linear theory for six field components that satisfies Coulomb's force law. Maxwell's equations are derived in covariant form from the appropriate action and solved by means of the retarded Greens function. The general formalism for the energy-momentum tensor of a field theory is introduced and applied to the electromagnetic field. From the Liénard-Wiechert potentials, Larmor's formula is derived in relativistic form, and the covariant expression for the Lorentz force is derived from the action. The main results are Maxwell's equations themselves, most compactly expressed in Lorenz gauge by the wave equation (1.100), the energy-momentum tensor (1.110) for the electromagnetic field, the Liénard-Wiechert potentials (1.117), the relativistic Larmor formula (1.138), its solid-angle integrated version (1.141) and its non-relativistic approximation (1.143), and finally the relativistic expression (1.147) for the Lorentz force.

1.3.1 Field Tensor and Sources

Electromagnetism is a classical field theory with six degrees of freedom, namely the three components each of the electric and magnetic fields $\vec{E}$ and $\vec{B}$. Fields are functions of space and time. Since special relativity teaches us that space and time are not independent, any field theory must explicitly be constructed to agree with the space-time structure of special relativity. The electromagnetic field must thus be expressed as a four-vector or a tensor field. Obviously, a four vector is not sufficient to describe six degrees of freedom. The simplest object available is a rank-2 tensor, which offers 16 independent components in its most general form. A symmetric rank-2 tensor in four dimensions still has ten independent components, while an antisymmetric rank-2 tensor has exactly the required six degrees of freedom. The simplest possibility to describe six degrees of freedom with a Lorentz-covariant object in four dimensions is thus provided by an antisymmetric field tensor F of rank two, whose components must satisfy

$$F^{\mu\nu} = -F^{\nu\mu} , \quad F_{\mu\nu} = -F_{\nu\mu} . \tag{1.77}$$

The antisymmetry is most conveniently ensured expressing the components of F as derivatives of a four-potential A with components

$$A^{\mu} = \begin{pmatrix} \Phi \\ \vec{A} \end{pmatrix} , \tag{1.78}$$

where Φ is the ordinary scalar potential and $\vec{A}$ is the three-dimensional vector potential. The components of the rank-$(2,0)$ field tensor are then written in the manifestly antisymmetric form

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \ . \tag{1.79}$$

They can be conveniently summarised as

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & \vec{E}^\top \\ -\vec{E} & \mathcal{B} \end{pmatrix} \tag{1.80}$$

where the matrix

$$\mathcal{B}_{ij} = \varepsilon_{ija} B^a \tag{1.81}$$

is formed from the components of the magnetic field. The fields themselves are thus given by

$$\vec{E} = -\frac{1}{c}\dot{\vec{A}} - \vec{\nabla}\Phi \ , \quad \vec{B} = \vec{\nabla} \times \vec{A} \ . \tag{1.82}$$

Caution As usual, ε_{ijk} is the totally antisymmetric Levi-Civita symbol, defined such that $\varepsilon_{ijk} = 0$ if any two of its indices are equal and ε_{ijk} is the signature of the permutation of the indices (ijk). ◀

Given our signature $(-,+,+,+)$ of the Minkowski metric, the associated rank-$(0,2)$ tensor has the components

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -\vec{E}^\top \\ \vec{E} & \mathcal{B} \end{pmatrix} \ . \tag{1.83}$$

?

Convince yourself that (1.80) and (1.83) are correct.

The source of the electromagnetic field is the four-current density j which has the components

$$(j^\mu) = \begin{pmatrix} \rho c \\ \vec{j} \end{pmatrix} \ , \tag{1.84}$$

where ρ is the charge density and $\vec{j}$ is the three-dimensional current density. Charge conservation is expressed by the vanishing four-divergence of the four-current,

$$\partial_\mu j^\mu = \frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \ . \tag{1.85}$$

1.3.2 Lorentz transform of the electromagnetic field

Changing from one inertial frame to another moving with a velocity $\vec{v} = c\vec{\beta}$ with respect to the original frame, the field tensor is Lorentz transformed according to

$$F'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta} \ . \tag{1.86}$$

Orienting both coordinate frames such that their $\hat{e}_z$ axes coincide with the direction of relative motion, the special Lorentz transform is represented by the matrix given in (1.26), and (1.86) gives the following transformation rules for the electric and magnetic field components:

$$E'_x = \gamma(E_x + \beta B_y) \ , \quad E'_y = \gamma(E_y - \beta B_x) \ , \quad E'_z = E_z \ ,$$
$$B'_x = \gamma(B_x - \beta E_y) \ , \quad B'_y = \gamma(B_y + \beta E_x) \ , \quad B'_z = B_z \ . \tag{1.87}$$

?

Confirm the transformation equations (1.87) for the electric- and magnetic-field components

While the field components in the direction of motion remain unchanged, the transverse components are enhanced by the Lorentz factor γ. In particular, a purely electric or magnetic field in one frame obtains a magnetic or electric component in the other, moving frame, repectively. It is, however, not possible to transform a purely electric field into a purely magnetic field or vice versa. This is easily understood because the Lorentz transform must keep all Lorentz invariants unchanged that can be formed from the field tensor. These invariants can be written as

$$F_{\mu\nu}F^{\mu\nu} = 2\left(\vec{B}^2 - \vec{E}^2\right) \ , \quad (*F)_{\mu\nu}F^{\mu\nu} = -4\vec{E}\cdot\vec{B} \ , \tag{1.88}$$

where $*F$ is the (Hodge-) dual field tensor. Any Lorentz transform must thus conserve $(\vec{E}^2 - \vec{B}^2)$ and $\vec{E}\cdot\vec{B}$. Starting with $\vec{B} = 0$ in one inertial frame first of all demands that $\vec{E}'$ and $\vec{B}'$ must remain perpendicular to each other in any inertial frame. By the invariance of $(\vec{E}^2 - \vec{B}^2)$, a complete conversion of a purely electric to a purely magnetic field would require

Caution The Hodge dual field tensor is obtained from the field tensor by replacing $\vec{E} \to \vec{B}$ and $\vec{B} \to -\vec{E}$,

$$(*F^{\mu\nu}) = \begin{pmatrix} 0 & \vec{B}^\top \\ -\vec{B} & -\mathcal{E} \end{pmatrix}$$

with $\mathcal{E}_{ij} = \varepsilon_{ija}E^a$. ◂

$$\vec{E}^2 = -\vec{B}'^2 \ , \tag{1.89}$$

which is only possible in the trivial case $\vec{E} = 0 = \vec{B}'$ because $\vec{E}^2$ and $\vec{B}'^2$ are positive definite otherwise.

?

Can you confirm Eqs. (1.88)?

One remark on the transformation formula (1.86) may be in order to avoid confusion. In Euclidean space, a transformation R from one coordinate frame to another changes the matrix representation of a tensor T according to

$$T' = RTR^{-1} = RTR^\top \tag{1.90}$$

if R is orthogonal, $R^{-1} = R^\top$. Although the matrix representation (1.26) of the Lorentz transform does not satisfy this relation, the Lorentz transform is still orthogonal in the sense that it leaves (Minkowski) scalar products invariant, just as orthogonal transformations in Euclidean space leave the Euclidean scalar product unchanged; see also the discussion of this issue in Sect. 1.1.2 above. For this reason, (1.86) remains valid for Lorentz transformations.

1.3.3 Maxwell's Equations

The dynamical equations of a field theory are the Euler-Lagrange equations applied to a Lagrange density which, for a linear theory like electrodynamics, must satisfy three conditions: It must be Lorentz invariant, it must contain at most quadratic terms in the field quantities to ensure a linear theory, and it must reproduce the Coulomb force law in the case of electrodynamics. The only Lagrangian that satisfies these criteria is

$$\mathcal{L} = \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} - \frac{1}{c}j_\mu A^\mu \ , \tag{1.91}$$

where the constants must be chosen such as to reproduce the measured coupling strength of the electromagnetic field to matter. The otherwise perfectly legitimate term $A_\mu A^\mu$ is excluded because it would give the electromagnetic field an effective mass and thus violate the Coulomb force law.

Since the field tensor depends on A^μ only through derivatives, it is invariant under the gauge transformation

$$A^\mu \to A^\mu + \partial^\mu\chi \ , \tag{1.92}$$

where χ is an arbitrary function of all four coordinates x^μ. At first sight, the Lagrangian (1.91) appears to violate gauge invariance, but Gauss' law applied to the action

$$S = \int \mathrm{d}^4x\,\mathcal{L} \tag{1.93}$$

shows that charge conservation (1.85) ensures gauge invariance.

Maxwell's equations are now the Euler-Lagrange equations

$$\partial^\nu \frac{\partial\mathcal{L}}{\partial(\partial^\nu A^\mu)} - \frac{\partial\mathcal{L}}{\partial A^\mu} = 0 \tag{1.94}$$

of the Lagrangian (1.91). They turn out to be

$$\partial_\nu F^{\mu\nu} = \frac{4\pi}{c} j^\mu \;, \tag{1.95}$$

which are four inhomogeneous equations. Since the field tensor is antisymmetric, it identically satisfies the equation

$$\partial_{[\alpha}F_{\beta\gamma]} = 0 \;, \tag{1.96}$$

which represents the homogeneous Maxwell equations. For $\alpha = 0$, $(\beta,\gamma) = (1,2)$, $(1,3)$ and $(2,3)$, the homogeneous equations (1.96) give

$$\boxed{\dot{\vec{B}} + c\vec{\nabla}\times\vec{E} = 0} \;, \tag{1.97}$$

while we find

$$\boxed{\vec{\nabla}\cdot\vec{B} = 0} \tag{1.98}$$

for $\alpha = 1$, $(\beta,\gamma) = (2,3)$. Setting $\mu = 0$ and $\mu = i$, the inhomogeneous equations (1.95) give

$$\boxed{\vec{\nabla}\cdot\vec{E} = 4\pi\rho \;, \quad c\vec{\nabla}\times\vec{B} - \dot{\vec{E}} = 4\pi\vec{j}} \;, \tag{1.99}$$

respectively.

With the definition (1.79) of the field tensor in terms of the four-potential and with the Lorenz gauge condition $\partial_\mu A^\mu = 0$, the inhomogeneous equations (1.95) can be cast into the form

$$\boxed{\Box A^\mu = -\frac{4\pi}{c} j^\mu} \;, \tag{1.100}$$

where $\Box = -\partial_0^2 + \vec{\nabla}^2$ is the d'Alembert operator. The particular solution of this inhomogeneous wave equation is given by the convolution of the source with the retarded Greens function

$$G(t-t', \vec{x}-\vec{x}') = \frac{1}{\left|\vec{x}-\vec{x}'\right|}\delta_\mathrm{D}\left(t - t' - \frac{\left|\vec{x}-\vec{x}'\right|}{c}\right) \;, \tag{1.101}$$

i.e. by

$$A^\mu(t,\vec{x}) = \frac{1}{c}\int \mathrm{d}^3x' \int \mathrm{d}t' G\left(t-t', \vec{x}-\vec{x}'\right)\, j^\mu\left(t', \vec{x}'\right) \;. \tag{1.102}$$

The Greens function (1.101) has an intuitive meaning (Figure 1.5). Its first factor, proportional to the inverse distance between the observer and the source,

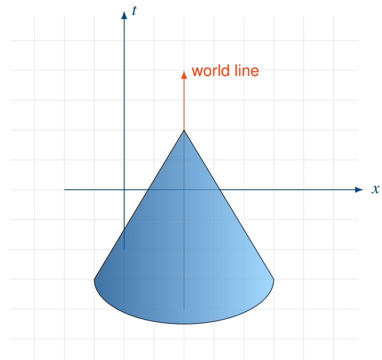

Figure 1.5 Illustration of the geometrical meaning of the retarded Green's function: All signals received from the observer on the red world line at a given instant of time must originate from the backward light cone ending at that time.

expresses Coulomb's force law, which is an immediate consequence of photons being massless. If photons had a mass, the Greens function would have a Yukawa shape with an exponential cut-off. The second factor, the delta function, shows that only such sources can influence the potential at the observer whose world lines intersect with the observer's backward light cone.

Since the Greens function is defined as

$$\Box G\left(t-t', \vec{x}-\vec{x}'\right) = -4\pi\delta_{\rm D}\left(t-t', \vec{x}-\vec{x}'\right) \;, \tag{1.103}$$

it represents any component of the four-potential created by a point source on the backward light cone of the observer. The convolution (1.102) assembles the complete four-potential by superposition of all contributing sources. This is possible only because electromagnetism is a linear field theory.

1.3.4 Energy-Momentum Conservation

A field theory with a Lagrangian $\mathcal{L}(q, \partial_\nu q)$ for a single field q and its derivatives $\partial_\nu q$ has the energy-momentum tensor

$$T^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu q)}\partial_\nu q - \mathcal{L}\delta^\mu_\nu \;, \tag{1.104}$$

which simply corresponds to the Legendre transformation leading from the Lagrange to the Hamilton function in classical mechanics. Should the expression (1.104) turn out to be asymmetric, it needs to be symmetrised to ensure the symmetry of the energy-momentum tensor. For the electromagnetic field, any A^γ can take the role of q, thus

$$T^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\gamma)}\partial_\nu A_\gamma - \mathcal{L}\delta^\mu_\nu \;. \tag{1.105}$$

With the Lagrange density of the free electromagnetic field,

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} \;, \tag{1.106}$$

this implies the energy-momentum tensor

$$T^{\mu}{}_{\nu} = \frac{1}{4\pi}\left(F^{\mu\lambda}F_{\nu\lambda} - \frac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\delta^{\mu}_{\nu}\right) \tag{1.107}$$

of the electromagnetic field. From the representations (1.80) and (1.83) of the field tensor, we find first

$$F^{\mu\lambda}F_{\nu\lambda} = \begin{pmatrix} -\vec{E}^2 & -\left(\vec{E}\times\vec{B}\right)^{\top} \\ \vec{E}\times\vec{B} & -E_iE_j + \delta_{ij}\vec{B}^2 - B_iB_j \end{pmatrix} \tag{1.108}$$

and confirm

$$F^{\alpha\beta}F_{\alpha\beta} = 2\left(\vec{B}^2 - \vec{E}^2\right) . \tag{1.109}$$

Thus, the energy-momentum tensor can be written as

$$T^{\mu}{}_{\nu} = \frac{1}{4\pi}\begin{pmatrix} -(\vec{E}^2+\vec{B}^2)/2 & \left(\vec{E}\times\vec{B}\right)^{\top} \\ -\vec{E}\times\vec{B} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \bar{T}_{ij} \end{pmatrix} , \tag{1.110}$$

where

$$\bar{T}_{ij} = \frac{1}{4\pi}\left[\left(\frac{1}{2}\vec{E}^2\delta_{ij} - E_iE_j\right) + \left(\frac{1}{2}\vec{B}^2\delta_{ij} - B_iB_j\right)\right] \tag{1.111}$$

are thew components of Maxwell's stress tensor, whose magnetic part will become important in magnetohydrodynamics. The energy density of the electromagnetic field is

$$\varepsilon = T_{00} = \frac{\vec{E}^2 + \vec{B}^2}{8\pi} . \tag{1.112}$$

?

Carry out all calculations leading to the results (1.110) and (1.111) for the energy-momentum tensor of the electromagnetic field.

The energy-momentum tensor satisfies the conservation equation

$$\partial_{\nu}T^{\mu\nu} = 0 \tag{1.113}$$

which, for $\mu = 0$, returns the continuity equation

$$\frac{\partial\varepsilon}{\partial t} + \vec{\nabla}\cdot\vec{S} = 0 \tag{1.114}$$

for the energy density, where the Poynting vector

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} \tag{1.115}$$

represents the energy-current density of the electromagnetic field.

1.3.5 Liénard-Wiechert Potentials and the Larmor Formula

A particle with charge q on a trajectory $\vec{r}_0(t)$ has the current density

$$j^{\mu} = q\begin{pmatrix} c \\ \vec{v} \end{pmatrix}\delta_{\mathrm{D}}\left[\vec{r} - \vec{r}_0(t)\right] . \tag{1.116}$$

When inserted into the convolution (1.102) with the retarded Greens function, this yields the Liénard-Wiechert potentials

$$\Phi(\vec{r}) = \frac{q}{R\left(1 - \hat{e}\cdot\vec{\beta}\right)} , \quad \vec{A}(\vec{r}) = \frac{q\vec{\beta}}{R\left(1 - \hat{e}\cdot\vec{\beta}\right)} = \Phi\vec{\beta} , \tag{1.117}$$

where the right-hand sides have to be evaluated at the retarded time

$$t' = t - \frac{R}{c} \;. \tag{1.118}$$

The vector $\vec{R} \equiv \vec{r} - \vec{r}_0(t')$ points from the retarded particle position to the observer, $R = |\vec{R}|$, and $\hat{e}$ is the unit vector in $\vec{R}$ direction,

$$\hat{e} = \frac{\vec{R}}{R} \;. \tag{1.119}$$

?

Can you confirm expressions (1.117) for the Liénard-Wiechert potentials of a point charge?

The fields $\vec{E}$ and $\vec{B}$ are obtained as the usual derivatives of Φ and $\vec{A}$, but it must be taken into account that the potentials are expressed in retarded coordinates, while we need the derivatives with respect to the observer's coordinates. The spatial derivatives of Φ are

$$\partial_i \Phi = -\frac{q}{\left(R - \vec{R}\cdot\vec{\beta}\right)^2}\left(\partial_i R - \beta_j \partial_i R_j - R_j \partial_i \beta_j\right) \;. \tag{1.120}$$

While the first two terms decrease $\propto R^{-2}$, the third decreases $\propto R^{-1}$. Aiming at the fields far away from any source, we retain only the latter, thus

$$(\partial_i \Phi)_{\rm far} = \frac{q R_j \partial_i \beta_j}{\left(R - \vec{R}\cdot\vec{\beta}\right)^2} = \frac{q\left(\hat{e}\cdot\dot{\vec{\beta}}\right)\partial_i t'}{R\left(1 - \hat{e}\cdot\vec{\beta}\right)^2} \;. \tag{1.121}$$

The remaining spatial derivative of the retarded time is

$$\partial_i t' = -\frac{\partial_i R}{c} = -\frac{R_j}{R}\partial_i R_j = -e_j\left(\frac{\delta_{ij}}{c} - \beta_j \partial_i t'\right) = -\frac{e_i}{c} + \left(\hat{e}\cdot\vec{\beta}\right)\partial_i t' \;. \tag{1.122}$$

This equation gives

$$\partial_i t' = -\frac{e_i}{c\left(1 - \hat{e}\cdot\vec{\beta}\right)} \;, \tag{1.123}$$

which implies with (1.121)

$$(\vec{\nabla}\Phi)_{\rm far} = -\frac{q\left(\hat{e}\cdot\dot{\vec{\beta}}\right)\hat{e}}{Rc\left(1 - \hat{e}\cdot\vec{\beta}\right)^3} \tag{1.124}$$

for the gradient of Φ in the far-field. The time derivative of $\vec{A}$ is

$$(\partial_t \vec{A})_{\rm far} = \Phi\partial_t\vec{\beta} + \vec{\beta}\partial_t\Phi = \frac{q\dot{\vec{\beta}}\partial_t t'}{R\left(1 - \hat{e}\cdot\vec{\beta}\right)} + \frac{q\vec{\beta}\left(\hat{e}\cdot\dot{\vec{\beta}}\right)\partial_t t'}{R\left(1 - \hat{e}\cdot\vec{\beta}\right)^2} \tag{1.125}$$

if we again drop all terms with a steeper R dependence than R^{-1} to isolate the far-field. Now, the time derivative of t' is given by

$$\partial_t t' = 1 - \frac{\partial_t R}{c} = 1 - \frac{R_j}{Rc}\partial_t R_j = 1 + \hat{e}\cdot\vec{\beta}\partial_t t' \;, \tag{1.126}$$

thus

$$\partial_t t' = \frac{1}{1 - \hat{e}\cdot\vec{\beta}} \;. \tag{1.127}$$

The far-field time derivative of $\vec{A}$ then turns into

$$\left(\partial_t \vec{A}\right)_{\text{far}} = \frac{q}{R\left(1-\hat{e}\cdot\vec{\beta}\right)^3}\left[\left(1-\hat{e}\cdot\vec{\beta}\right)\dot{\vec{\beta}}+\vec{\beta}\left(\hat{e}\cdot\dot{\vec{\beta}}\right)\right] . \tag{1.128}$$

From this, together with (1.124), and using the identity $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\cdot\vec{c})\vec{b} - (\vec{a}\cdot\vec{b})\vec{c}$ twice, we find the electric field far from the source,

$$\vec{E}_{\text{far}} = \frac{q}{Rc\left(1-\hat{e}\cdot\vec{\beta}\right)^3}\hat{e}\times\left[\left(\hat{e}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right] . \tag{1.129}$$

The magnetic field is

$$\vec{B} = \vec{\nabla}\times\vec{A} = \Phi\vec{\nabla}\times\vec{\beta} - \vec{\beta}\times\vec{\nabla}\Phi . \tag{1.130}$$

Taking the curl of the velocity $\vec{\beta}$, we must be aware that $\vec{\beta}$ depends on position through the retarded time t'. In components, we have

$$\left(\vec{\nabla}\times\vec{\beta}\right)_i = \varepsilon_{ijk}\partial_j\beta_k = \varepsilon_{ijk}\dot{\beta}_k\partial_j t' . \tag{1.131}$$

With the help of (1.123), we then find

$$\vec{\nabla}\times\vec{\beta} = -\frac{\hat{e}\times\dot{\vec{\beta}}}{c\left(1-\hat{e}\cdot\vec{\beta}\right)} , \tag{1.132}$$

which, together with (1.124), allows us to write

$$\vec{B}_{\text{far}} = -\frac{q}{Rc\left(1-\hat{e}\cdot\vec{\beta}\right)^3}\hat{e}\times\left[\dot{\vec{\beta}}+\hat{e}\times\left(\vec{\beta}\times\dot{\vec{\beta}}\right)\right] . \tag{1.133}$$

> **?** Convince yourself by your own calculation that expressions (1.129) and (1.133) for the electric and magnetic fields far from the source are correct.

Comparing to (1.129), it is straightforward to confirm that

$$\vec{B}_{\text{far}} = \hat{e}\times\vec{E}_{\text{far}} . \tag{1.134}$$

Using this result, the Poynting vector far away from the source is

$$\vec{S} = \frac{q^2}{4\pi R^2 c\left(1-\hat{e}\cdot\vec{\beta}\right)^6}\left|\hat{e}\times\left[\left(\hat{e}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]\right|^2\hat{e} . \tag{1.135}$$

This quantifies the energy received per unit area per unit time by the observer. We now need to distinguish between a time interval $\mathrm{d}t$ measured by the observer and the corresponding interval $\mathrm{d}t'$ of the retarded time. The latter is the time interval during which the source needs to emit for the observer to see its radiation for the time interval $\mathrm{d}t$. Since, according to (1.127), the retarded time interval $\mathrm{d}t'$ is related to the time interval $\mathrm{d}t$ measured by the observer through

$$\mathrm{d}t = \left(1-\hat{e}\cdot\vec{\beta}\right)\mathrm{d}t' , \tag{1.136}$$

the energy emitted per the observer's unit time $\mathrm{d}t$ into the solid angle element $\mathrm{d}\Omega$ is

$$\mathrm{d}E = \vec{S}\cdot\hat{e}R^2\mathrm{d}\Omega\mathrm{d}t = \frac{q^2}{4\pi c\left(1-\hat{e}\cdot\vec{\beta}\right)^5}\left|\hat{e}\times\left[\left(\hat{e}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]\right|^2\mathrm{d}\Omega\mathrm{d}t' , \tag{1.137}$$

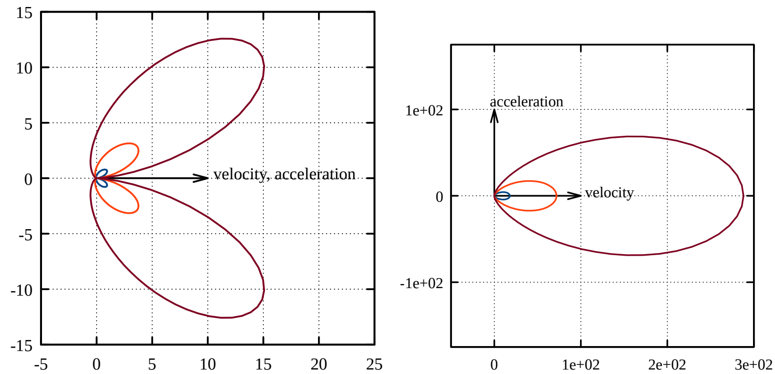

Figure 1.6 The radiation power according to the relativistic Larmor formula is illustrated for a charge accelerated parallel (left) and perpendicular (right) to its relativistic velocity. Three curves are given for three arbitrary values of the acceleration. Notice the different scales of the two plots!

and thus the power emitted per unit solid angle and per unit retarded time $\mathrm{d}t'$ is

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{q^2}{4\pi c\left(1 - \hat{e}\cdot\vec{\beta}\right)^5}\left|\hat{e}\times\left[\left(\hat{e}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]\right|^2 . \tag{1.138}$$

This is the relativistic Larmor formula which describes the power radiated by a source per unit solid angle (Figures 1.6 and 1.7).

The total emitted power is the solid-angle integral of (1.138). This calculation is not difficult to carry out, but lengthy. Perhaps the most straightforward way begins by expanding the double vector product using the identity

$$\vec{a}\times\left(\vec{b}\times\vec{c}\right) = (\vec{a}\cdot\vec{c})\,\vec{b} - \left(\vec{a}\cdot\vec{b}\right)\vec{c}\,, \tag{1.139}$$

followed by squaring the result. Then, it is useful to introduce coordinates such that the velocity $\vec{\beta}$ points into the $\hat{e}_z$ direction, $\vec{\beta} = \beta\hat{e}_z$, the acceleration $\dot{\vec{\beta}}$ falls into the x-z plane, $\dot{\vec{\beta}} = \dot{\beta}(\sin\alpha\hat{e}_x + \cos\alpha\hat{e}_z)$, and

$$\hat{e} = \begin{pmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{pmatrix} . \tag{1.140}$$

Then, the ϕ and θ integrations can be carried out in this order, giving the result

$$P = \frac{2e^2}{3c}\gamma^6\left[\dot{\vec{\beta}}^2 - \left(\vec{\beta}\times\dot{\vec{\beta}}\right)^2\right] . \tag{1.141}$$

?

Can you confirm that integrating the Larmor formula (1.138) over the solid angle results in (1.141)?

The factor γ^6 is most remarkable: A relativistically moving charge with a high Lorentz factor radiates with an enormous power. For non-relativistically moving charges, equations (1.138) and (1.141) simplify to

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{q^2}{4\pi c}\left|\hat{e}\times\left(\hat{e}\times\dot{\vec{\beta}}\right)\right|^2 = \frac{q^2}{4\pi c}\left|\dot{\vec{\beta}} - \left(\dot{\vec{\beta}}\cdot\hat{e}\right)\hat{e}\right|^2 = \frac{q^2}{4\pi c}\dot{\beta}_\perp^2\,, \tag{1.142}$$

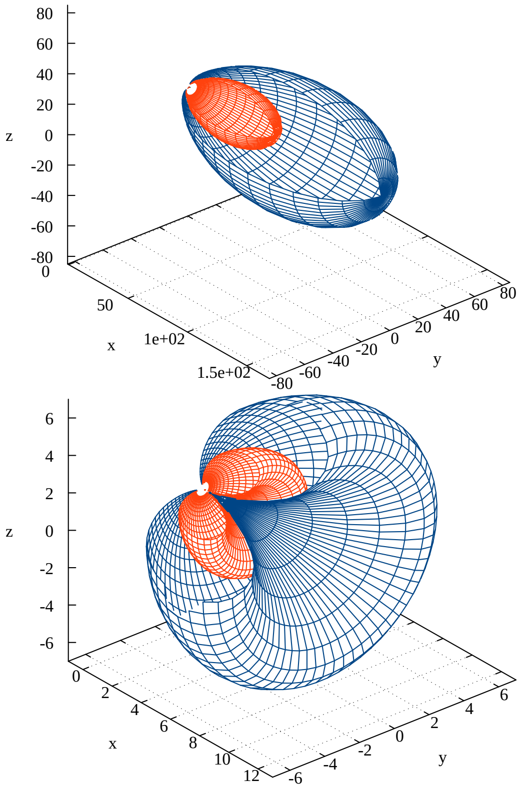

Figure 1.7 Three-dimensional illustrations of the radiation power of two accelerated charges with $\beta = 0.5$. A charge accelerated perpendicular to its direction of motion has its emission peaked strongly into the forward direction of its motion (top panel), while the radiation of a charge accelerated parallel to its direction of motion is emitted into a collar surrounding its trajectory (bottom panel).

where $\dot{\vec{\beta}}_\perp$ is the acceleration perpendicular to $\hat{e}$, and

$$P = \frac{2q^2}{3c}\dot{\beta}^2 \ . \tag{1.143}$$

1.3.6 The Lorentz Force

The action for a relativistic particle with mass m and charge q in an electromagnetic field with vector potential A^μ is

$$S = -mc^2 \int \mathrm{d}\tau + \frac{q}{c} \int A_\mu \mathrm{d}x^\mu \ . \tag{1.144}$$

This is the simplest Lorentz-invariant expression that can be formed from the only Lorentz-invariant quantity of a free particle, i.e. its proper time τ, the four potential A^μ and the coordinates x^μ of the particle trajectory. Variation of the action (1.144) with respect to the particle trajectory under a fixed four-potential A^μ and equating the result to zero leads to the equation of motion

$$m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{q}{c}F^\mu{}_\nu u^\nu \; . \tag{1.145}$$

With $u^0 = \gamma c$, $u^i = \gamma v^i$ and $\mathrm{d}\tau = \gamma^{-1}\mathrm{d}t$, the 0-component of this equation means

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma mc^2\right) = q\vec{E}\cdot\vec{v}\,, \tag{1.146}$$

showing that the work done by the electric field changes the energy γmc^2 of the particle. The spatial components give

$$m\frac{\mathrm{d}\left(\gamma\vec{v}\right)}{\mathrm{d}t} = q\vec{E} + \frac{q}{c}\vec{v}\times\vec{B} \; . \tag{1.147}$$

For non-relativistic motion, $\gamma = 1$, and (1.147) reproduces the common equation of motion under the Lorentz force.

?

Derive the equation of motion (1.145) as the Euler-Lagrange equation of the action (1.144).

Problems

1. Starting from the Lorentz transform (1.86) of the electromagnetic field tensor, expressly derive the Lorentz transform (1.87) of the field components.

2. Show by explicit calculation that Maxwell's equation in three-dimensional form follow from their relativistic forms (1.95) and (1.96).

3. Convince yourself that the components of the energy-momentum tensor (1.110) have the appropriate physical units.

4. From the invariants (1.88) of the electromagnetic field tensor, derive the following statements:

   (a) If $\vec{E}$ and $\vec{B}$ have the same amplitude $|\vec{E}| = |\vec{B}|$ in one inertial frame, then also in all other inertial frames.
   (b) If $\vec{E}$ and $\vec{B}$ are orthogonal in one inertial frame, then also in all other inertial frames.

5. Apply the Larmor formula to the classical picture of an electron in a hydrogen atom.

   (a) Decide whether the non-relativistic approximation of the Larmor formula can be applied.
   (b) Estimate the classical lifetime of a hydrogen atom.

6. Derive the electromagnetic field of a point charge q uniformly moving with the velocity $\vec{v}_0$.

(a) Calculate the Liénard-Wiechert potentials (1.117) for a point charge moving with constant velocity along a straight line and compute the electromagnetic fields from them. *Hint*: Since the velocity is constant, we never need the retarded time itself, but only the separation R between the charge and any point $\vec{x}$ at the retarded time. Introducing the vector $\vec{\omega} = \vec{x} - \vec{v}_0 t$ helps greatly.

(b) Find the fields by a suitable Lorentz transform and compare the two results.

1.4 Elementary kinetic theory

This section serves a dual purpose. The discussion of the Boltzmann equation and the BBGKY hierarchy prepares the derivation of the hydrodynamical equations later in this book. The Fokker-Planck equation derived thereafter from a diffusion approximation of the collision terms occurs under a variety of circumstances in astrophysics, from radiation transport to stellar dynamics. The main results are the Boltzmann equation (1.156), the master equation (1.161), the Fokker-Planck equation in its original form (1.163), its form (1.172) with one of the diffusion coefficients eliminated by equilibrium considerations, and its form (1.180) for small changes in absolute momentum.

1.4.1 The BBGKY hierarchy and the Boltzmann equation

Kinetic theory describes how *ensembles* of particles change in time, in absence or in presence of mutual collisions. In classical mechanics, generalised coordinates q_i are assigned to the degrees of freedom that the system under consideration has. The number of degrees of freedom d depends on the number of components of the system and their mutual relations to each other. If the system consists of N independent point particles in three-dimensional space, $d = 3N$. If those particles are linked to form a solid body, $d = 6$, because only three degrees of translational and three degrees of rotational freedom remain. By Newton's second law, two initial conditions must be given for each degree of freedom, which can be chosen to be the generalised coordinates q_i and the associated velocities, $\dot{q}_i$, at some initial time.

If the system can be described by a Lagrange function $L(q_i, \dot{q}_i, t)$, the canonically conjugated momenta

$$p_i = \frac{\partial L(q_i, \dot{q}_i, t)}{\partial \dot{q}_i} \tag{1.148}$$

can be substituted for the velocities $\dot{q}_i$ by the Legendre transform

$$H(q_i, p_i, t) = \sum_{i=1}^{d} \dot{q}_i p_i - L\left[q_i, \dot{q}_i(p_i), t\right] \ , \tag{1.149}$$

leading to the Hamilton function $H(q_i, p_i, t)$. The equations of motion for all degrees of freedom are then Hamilton's equations,

$$\dot{q}_i = \frac{\partial H(q_i, p_i, t)}{\partial p_i} \ , \quad \dot{p}_i = -\frac{\partial H(q_i, p_i, t)}{\partial q_i} \ . \tag{1.150}$$

The physical state of such a system is fully characterised by the d generalised coordinates $\vec{q} = (q_1, \ldots, q_d)$ and their d canonically conjugated momenta $\vec{p} = (p_1, \ldots, p_d)$. The generalised coordinates $\vec{q}$ span the configuration space of the system. Together with their conjugate momenta, they span the $2d$-dimensional phase space.

It is by no means unique how the $2d$ phase-space coordinates are to be divided into generalised coordinates and their conjugate momenta. Canonical transformations applied to phase space leave Hamilton's equations invariant, but can turn coordinates into momenta and vice versa.

The classical physical state of the system is given by the system's location in the $2d$-dimensional phase space. Statistical mechanics is not interested in the phase-space coordinates of all particles in an ensemble. Rather, it divides phase space into cells of small, but finite size, sums the number of particles in each cell and studies the time evolution of this number instead of the time evolution of each individual pair (q_i, p_i) of phase-space coordinates. We thus introduce a distribution function $f^{(d)}(t, \vec{q}, \vec{p}\,)$ such that the probability for finding the system in a small phase-space cell around to the phase-space point $(\vec{q}, \vec{p}\,)$ at time t is

$$\mathrm{d}P^{(d)}\,(t, \vec{q}, \vec{p}\,) = f^{(d)}\,(t, \vec{q}, \vec{p}\,)\;\mathrm{d}^d q\,\mathrm{d}^d p\;. \tag{1.151}$$

For systems with very many degrees of freedom, the full phase-space distribution function $f^{(d)}$ becomes utterly unmanageable, apart from the fact that the complete knowledge of the evolution of all d degrees of freedom is then neither desired nor necessary. Rather, we are then interested in the reduced phase-space distribution function $f^{(k)}$, obtained by integrating $f^{(d)}$ over $d - k$ coordinates and momenta,

$$f^{(k)}(t, q_1, \ldots, q_k, p_1, \ldots p_k) = \int \mathrm{d}q_{k+1} \ldots \mathrm{d}q_d \int \mathrm{d}p_{k+1} \ldots \mathrm{d}p_d\, f^{(d)}(t, \vec{q}, \vec{p}\,)\;. \tag{1.152}$$

By Liouville's theorem and Hamilton's equations, the time evolution of the full phase-space distribution function $f^{(d)}$ is determined by Liouville's equation

$$\frac{\partial f^{(d)}}{\partial t} + \dot{q}_i \frac{\partial f^{(d)}}{\partial q_i} + \dot{p}_j \frac{\partial f^{(d)}}{\partial p_j} = \frac{\partial f^{(d)}}{\partial t} + \frac{\partial H}{\partial p_i}\frac{\partial f^{(d)}}{\partial q_i} - \frac{\partial H}{\partial q_j}\frac{\partial f^{(d)}}{\partial p_j} = 0\;. \tag{1.153}$$

Searching for an evolution equation for any of the reduced phase-space distribution functions $f^{(k)}$, we have to integrate Liouville's equation over $d - k$ degrees of freedom and sort terms accordingly. It then appears that the evolution of the reduced distribution function $f^{(k)}$ depends on the reduced distribution function at the next higher level, $f^{(k+1)}$. This establishes the so-called BBGKY hierarchy of equations of motion for the reduced distribution functions, where the acronym stands for the authors Born, Bogoliubov, Green, Kirkwood and Yvon.

To see what the BBGKY hierarchy means, let us begin with the reduced phase-space distribution $f^{(1)}$ for a single degree of freedom. It will depend on the distribution function $f^{(2)}$ for two degrees of freedom, which expresses the notion that individual degrees of freedom do not evolve in isolation, but in correlation with others. In an ensemble of particles, the motion of a single particle is determined by two-body correlations with other particles, which in turn are affected by three-body correlations, and so forth. Clearly, the BBGKY hierarchy needs to be terminated somewhere, or closed, for us to make any progress. This closure is typically set by ignoring any correlations higher than a certain order.

We are particularly interested in the evolution of the distribution function for *single particles*. Let us therefore imagine that we have an ensemble of N point particles with $d = 3N$ degrees of freedom. We then integrate out all $3N - 3 = 3(N - 1)$ degrees of freedom belonging to $N - 1$ of the N particles and arrive at an evolution equation for the one-particle distribution function. According to the BBGKY hierarchy, this evolution equation will contain two-particle correlations. Closure can now be achieved by assuming that any two particles are *statistically uncorrelated*. The joint probability for finding a pair of particles at two positions in phase space is then simply the product of the probabilities for finding one of the particles at one position and the other at the other position. The two-particle distribution function can then be written as a product of one-particle distribution functions.

This closure condition means that any two particles affect each other's motion exclusively by direct two-body collisions. They move independently until they collide, and continue moving independently after the collision. This is possible if the interaction potential between any two particles is short-ranged compared to the mean inter-particle distance.

Following these considerations, we introduce a one-particle distribution function $f(t, \vec{q}, \vec{p}\,)$ by integrating $f^{(d)}$ over all but those degrees of freedom that belong to a single particle. For an ensemble of point particles in three-dimensional space, $f(t, \vec{q}, \vec{p}\,)$ is then defined on an effective, six-dimensional phase space. Moreover, we normalize the distribution $f(t, \vec{q}, \vec{p}\,)$ such that

$$f(t, \vec{q}, \vec{p}\,)\, \mathrm{d}^3q\, \mathrm{d}^3p = \mathrm{d}N \tag{1.154}$$

is the number of particles expected to be found within the infinitesimal phase-space volume $\mathrm{d}\Gamma = \mathrm{d}^3q\mathrm{d}^3p$ around the phase-space position $(\vec{q}, \vec{p}\,)$. For this one-particle phase-space distribution function $f(t, \vec{q}, \vec{p}\,)$, Liouville's equation reduces to Boltzmann's equation,

$$\frac{\partial f}{\partial t} + \dot{\vec{q}} \cdot \frac{\partial f}{\partial \vec{q}} + \dot{\vec{p}} \cdot \frac{\partial f}{\partial \vec{p}} = C[f]\ , \tag{1.155}$$

where the term $C[f]$ is called *collision term*: According to our closure condition for the BBGKY hierarchy, particle interactions are determined by direct particle collisions only and thus by the one-particle distribution function itself. The collision term must then be a functional of f. For a Hamiltonian system with Hamilton function $H = H(t, \vec{q}, \vec{p}\,)$, Boltzmann's equation reads

$$\frac{\partial f}{\partial t} + \frac{\partial H}{\partial \vec{p}} \cdot \frac{\partial f}{\partial \vec{q}} - \frac{\partial H}{\partial \vec{q}} \cdot \frac{\partial f}{\partial \vec{p}} = C[f]\ . \tag{1.156}$$

If, as usual, the Hamilton function can be written as $H = T + V$, with the kinetic energy T depending on the conjugate momenta $\vec{p}$ only and a potential energy V depending only on the generalised coordinates $\vec{q}$, and if further $T = \vec{p}^{\,2}/(2m)$, then we can write

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial \vec{q}} \cdot \frac{\vec{p}}{m} - \frac{\partial f}{\partial \vec{p}} \cdot \frac{\partial V}{\partial \vec{q}} = C[f]\ . \tag{1.157}$$

1.4.2 Collision terms

In presence of collisions, the phase-space density changes schematically according to

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \text{gain} - \text{loss} \ , \tag{1.158}$$

where the gain and loss terms are due to scattering into and out of the phase-space element $\mathrm{d}\vec{w}$ under consideration. Let $\psi(\vec{w}, \delta\vec{w})\mathrm{d}\delta\vec{w}\mathrm{d}t$ be the transition probability due to scattering by an amount $\delta\vec{w}$ from $\vec{w}$ to $\vec{w} + \delta\vec{w}$ within the time interval $\mathrm{d}t$. Typically, ψ would be quantified by a scattering cross section. Then, the gain term is

$$\text{gain} = \int \mathrm{d}\delta\vec{w}\, \psi(\vec{w} - \delta\vec{w}, \delta\vec{w}) f(t, \vec{w} - \delta\vec{w}) \tag{1.159}$$

since the integral quantifies the expected number of particles moving per unit time from the phase-space coordinates $\vec{w} - \delta\vec{w}$ to the phase-space coordinates $\vec{w}$: It multiplies the number of particles at the original phase-space point with their transition probability per unit time and integrates over all possible changes $\delta\vec{w}$. Similarly, the loss term is

$$\text{loss} = \int \mathrm{d}\delta\vec{w}\, \psi(\vec{w}, \delta\vec{w}) f(t, \vec{w}) \ . \tag{1.160}$$

Inserting these gain and loss terms (1.159) and (1.160) into (1.158) yields the so-called master equation

$$\frac{\mathrm{d}f(t, \vec{w})}{\mathrm{d}t} = \int \mathrm{d}\delta\vec{w} \left[\psi(\vec{w} - \delta\vec{w}, \delta\vec{w}) f(t, \vec{w} - \delta\vec{w}) - \psi(\vec{w}, \delta\vec{w}) f(t, \vec{w}) \right] \ , \tag{1.161}$$

describing the change of the phase-space density due to the collisions causing the transition probability ψ in phase space.

1.4.3 Diffusion in phase space: The Fokker-Planck approximation

We study the time evolution of the phase-space density f here under the quite relevant assumption that the phase-space coordinates of particles change only by small amounts in individual collisions. Then, the particles diffuse in phase space and their phase-space density changes gradually in a way that can be described with two diffusion coefficients. As we shall see in the course of this treatment, it is sufficient for this approximation if the absolute values of the phase-space coordinates change only very little in each collision, while the scattering angles can even be large. Under these circumstances, this diffusion approximation is most useful to describe all kinds of particle ensembles which either have low mass or low energy and interact with another particle ensemble of high mass or high energy. The equation describing how the phase-space density f changes with time under this approximation is called the Fokker-Planck equation. Its derivation, and general methods for its solution, are the main subject of the following treatment.

Specifically, let us assume that conditions are such that it is permissible to assume that the change $\Delta\vec{w}$ in the phase-space coordinates is small enough for the transition probability and the phase-space density at $\vec{w} - \Delta\vec{w}$ to be approximated by Taylor expansions up to second order,

$$\begin{aligned}\psi(\vec{w} - \delta\vec{w}, \delta\vec{w})f(t, \vec{w} - \delta\vec{w}) \approx\ & \psi(\vec{w}, \delta\vec{w})f(t, \vec{w}) \\ & - \frac{\partial}{\partial w_i}\left[\psi(\vec{w}, \delta\vec{w})f(t, \vec{w})\right]\delta w_i \\ & + \frac{1}{2}\frac{\partial^2}{\partial w_i \partial w_j}\left[\psi(\vec{w}, \delta\vec{w})f(t, \vec{w})\right]\delta w_i \delta w_j\ .\end{aligned} \tag{1.162}$$

Inserting this into the master equation (1.161) leads us already to the Fokker-Planck equation

$$\frac{\mathrm{d}f(\vec{w})}{\mathrm{d}t} = -\frac{\partial}{\partial w_i}\left[f(\vec{w})\,D_1^i(\vec{w})\right] + \frac{\partial^2}{\partial w_i \partial w_j}\left[f(\vec{w})\,D_2^{ij}(\vec{w})\right]\ , \tag{1.163}$$

which approximates scattering as a second-order diffusion process in phase space. The first- and second-order diffusion coefficients are

$$\begin{aligned}D_1^i(\vec{w}) &= \int \mathrm{d}\delta\vec{w}\,\psi(\vec{w}, \delta\vec{w})\,\delta w_i\ , \\ D_2^{ij}(\vec{w}) &= \frac{1}{2}\int \mathrm{d}\delta\vec{w}\,\psi(\vec{w}, \delta\vec{w})\,\delta w_i \delta w_j\ .\end{aligned} \tag{1.164}$$

The first-order coefficient D_1^i integrates the change δw_i in the phase-space coordinate w_i over the transition probability per unit time and thus quantifies the mean change of w_i per unit time. Similarly, the second-order coefficient D_2^{ij} quantifies the variances D_2^{ii} of the changes in w_i, and the covariances D_2^{ij} of different phase-space coordinates w_i and w_j for $i \neq j$. Thus, the combined vector with components D_1^i is the mean change per unit time of the position vector $\vec{w}$ in phase space, while D_2^{ij} is the covariance matrix of all individual changes.

?

Verify the Fokker-Planck equation (1.163) and the expressions (1.164) for the diffusion coefficients by your own derivation.

Suppose now that any change in the spatial coordinates is irrelevant, for example because all relevant particle species are homogeneously distributed in space. In fact, this assumption is much less restrictive than it might seem. It can also be satisfied statistically in the sense that although particles may move in space, the number of particles moving away from a specific point in space is compensated by an equal number moving there. In other words, what we set out to consider now is a dynamical spatial equilibrium. Then, we can concentrate on the d-dimensional momentum subspace of phase space, restrict $\vec{w} = \vec{p}$ and $\delta\vec{w} = \delta\vec{p}$ and consider the phase-space distribution function f as a function of $(t, \vec{p})$ only. The total time derivative of $f(t, \vec{p})$ then equals its partial time derivative, because

$$\frac{\partial f(t, \vec{p})}{\partial \vec{q}} = 0 \quad \text{and} \quad \frac{\partial f(t, \vec{p})}{\partial \vec{p}} \cdot \dot{\vec{p}} = 0 \quad \text{for} \quad \dot{\vec{p}} = -\frac{\partial H}{\partial \vec{q}} = 0\ . \tag{1.165}$$

Then, the Fokker-Planck equation (1.163) simplifies to a partial differential equation in time and momentum only,

$$\begin{aligned}\frac{\partial f(t, \vec{p})}{\partial t} &= -\frac{\partial}{\partial p_i}\left[f(t, \vec{p})D_1^i(\vec{p})\right] + \frac{\partial^2}{\partial p_i \partial p_j}\left[f(t, \vec{p})D_2^{ij}(\vec{p})\right] \\ &= -\frac{\partial}{\partial p_i}\left[\left(D_1^i(\vec{p}) - \frac{\partial}{\partial p_j}D_2^{ij}(\vec{p})\right)f(t, \vec{p}) - D_2^{ij}(\vec{p})\frac{\partial f(t, \vec{p})}{\partial p_j}\right]\ .\end{aligned} \tag{1.166}$$

This equation manifestly has the form of a continuity equation, where the term in brackets represents the current density $\vec{j}_p$ in momentum space,

$$\frac{\partial f(t,\vec{p})}{\partial t} + \vec{\nabla}_p \cdot \vec{j}_p = 0 \,, \tag{1.167}$$
$$j_p^i = \left(D_1^i(\vec{p}) - \frac{\partial}{\partial p_j} D_2^{ij}(\vec{p}) \right) f(t,\vec{p}) - D_2^{ij}(\vec{p}) \frac{\partial f(t,\vec{p})}{\partial p_j} \,.$$

At this point, it is important to note that the two diffusion coefficients D_1^i and D_2^{ij} are generally not independent. In an equilibrium situation, the momentum current $\vec{j}_p$ must vanish. Setting the components $j_p^i = 0$ in (1.167) for an equilibrium phase-space distribution $\bar{f}(t,\vec{p})$ implies that then the coefficient D_1^i can be expressed by D_2^{ij} and the derivative of $\bar{f}(t,\vec{p})$ with respect to the momentum,

$$D_1^i(\vec{p}) = \frac{\partial D_2^{ij}(\vec{p})}{\partial p_j} + D_2^{ij}(\vec{p}) \frac{\partial \ln \bar{f}(t,\vec{p})}{\partial p_j} \,. \tag{1.168}$$

However, since both coefficients do not depend on the specific form of f, we can now use them in the more general situation of an arbitrary phase-space distribution. Inserting the relation (1.168) into (1.167), the derivative of D_2^{ij} with respect to the momenta cancels, and the momentum current

$$j_p^i = -D_2^{ij}(\vec{p}) f(t,\vec{p}) \frac{\partial}{\partial p_j} \left[\ln f(t,\vec{p}) - \ln \bar{f}(t,\vec{p}) \right] \tag{1.169}$$

is shown to be driven by the momentum gradient of the ratio between the actual and the equilibrium phase-space distributions.

Example: Maxwellian momentum distribution

Suppose, for example, that the equilibrium distribution of the particle species under consideration can be described as a Maxwellian momentum distribution with a temperature $\bar{T}$. Then,

$$\bar{f}(t,\vec{p}) \propto \exp\left(-\frac{p^2}{2mk\bar{T}} \right) \,, \qquad \frac{\partial \ln \bar{f}(t,\vec{p})}{\partial p_j} = -\frac{p_j}{mk\bar{T}} \,, \tag{1.170}$$

the components of the momentum current simplify to

$$j_p^i = -D_2^{ij}(\vec{p}) f(t,\vec{p}) \left[\frac{\partial \ln f(t,\vec{p})}{\partial p_j} + \frac{p_j}{mk\bar{T}} \right] \,, \tag{1.171}$$

and the Fokker-Planck equation becomes

$$\frac{\partial f(t,\vec{p})}{\partial t} - \frac{\partial}{\partial p_i} \left[D_2^{ij}(\vec{p}) f(t,\vec{p}) \left(\frac{\partial f(t,\vec{p})}{\partial p_j} + \frac{p_j}{mk\bar{T}} \right) \right] = 0 \,. \tag{1.172}$$

◂

1.4.4 Diffusion in absolute momentum

Quite frequently, the scattering process changes the absolute value of the momentum by a small amount only, while the scattering angle may be large. Then, the diffusion approximation is still valid in terms of the absolute momentum, but not in the full three-dimensional momentum space any more. In other words, momentum can then be considered as slowly diffusing between spherical shells in momentum space, while its direction angles may be vastly redistributed from one shell to another. Instead of the phase-space density $f(t,\vec{p})$, we must then consider the density $f(t,p)p^2$ of particles in absolute momentum, irrespective of its direction. The Fokker-Planck approximation then still applies between momentum shells, and the Fokker-Planck equation becomes

$$\frac{\partial\left(fp^2\right)}{\partial t} = \frac{\partial}{\partial p}\left[\left(D_1 + \frac{\partial D_2}{\partial p}\right)\left(fp^2\right) + D_2\frac{\partial\left(fp^2\right)}{\partial p}\right] , \tag{1.173}$$

with the diffusion coefficients

$$D_1(p) = \int \mathrm{d}\delta p\, \psi(p,\delta p)\delta p\ , \quad D_2(p) = \frac{1}{2}\int \mathrm{d}\delta p\, \psi(p,\delta p)\delta p^2\ . \tag{1.174}$$

Both coefficients are now one-dimensional. The first-order coefficient D_1 is the mean momentum change per unit time, while the second-order coefficient D_2 is its mean-square.

We can now express the Fokker-Planck equation as a radial diffusion equation in momentum space,

$$\frac{\partial f}{\partial t} + \frac{1}{p^2}\frac{\partial\left(j_p p^2\right)}{\partial p} = 0\ , \quad j_p = \left(D_1 + \frac{\partial D_2}{\partial p}\right)f + \frac{D_2}{p^2}\frac{\partial\left(fp^2\right)}{\partial p}\ , \tag{1.175}$$

where now j_p is the radial component of the momentum current. Notice that the operator applied to the momentum current is the divergence in spherical polar coordinates, so the meaning of the equation has not changed: It remains a conservation equation, expressing that any change in phase-space density is caused by a momentum current.

Again, j_p must vanish in an equilibrium situation, expressed by an equilibrium phase-space density $\bar{f}$. This requirement establishes the relation

$$D_1 = -\left(\frac{2D_2}{p} + \frac{\partial D_2}{\partial p}\right) - D_2\frac{\partial \ln \bar{f}}{\partial p} \tag{1.176}$$

between D_1 and D_2. Inserting this result into the current density in (1.175) gives, after some straightforward rearrangement,

$$j_p = D_2 f\frac{\partial}{\partial p}\left(\ln f - \ln \bar{f}\right) = D_2 f\frac{\partial}{\partial p}\ln\frac{f}{\bar{f}}\ . \tag{1.177}$$

?

Convince yourself of the relations (1.176) and (1.177) under the conditions discussed here.

1.4.5 Calculation of the diffusion coefficient D_2

For an actual calculation of the diffusion coefficient D_2, we return to its definition in (1.164) or the more specialised form (1.174) and recall that the physical meaning of D_2 is (one half) the mean-squared momentum change per unit time of the particle species considered,

$$D_2 = \frac{1}{2}\left\langle \delta p^2\right\rangle\ . \tag{1.181}$$

Example: Maxwellians with different temperatures

To give an example, let us assume that both the actual and the equilibrium phase-space distributions, f and $\bar{f}$, are Maxwellians characterised by two different temperatures T and $\bar{T}$, respectively. Then,

$$\frac{\partial \ln f}{\partial p} - \frac{\partial \ln \bar{f}}{\partial p} = -\frac{p}{mkT}\left(1 - \frac{T}{\bar{T}}\right) , \tag{1.178}$$

the momentum current density becomes

$$j_p = -D_2 \frac{pf}{mkT}\left(1 - \frac{T}{\bar{T}}\right) , \tag{1.179}$$

and the Fokker-Planck equation reduces to

$$\frac{\partial f}{\partial t} = \frac{1}{p^2 mkT}\left(1 - \frac{T}{\bar{T}}\right)\frac{\partial}{\partial p}\left(D_2 p^3 f\right) . \tag{1.180}$$

◀

To illustrate this, consider a species of heavy particles with mass M embedded in a sea of light particles with mass $m \ll M$. Then, the energy of the heavy particles is almost unchanged by the collisions with the light particles, while momentum conservation implies a small change δp in absolute momentum determined by

$$\delta p^2 = 2q^2(1 - \cos\theta) \tag{1.182}$$

per collision, if q and θ are the momentum and the scattering angle of the light particle. The probability of a light particle with velocity $v = q/m$ scattering off a heavy particle per unit time into the solid-angle element $\mathrm{d}\Omega$ is

$$nv\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = \frac{nq}{m}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega , \tag{1.183}$$

where n is the number density of light particles. Thus, the mean-squared momentum change per unit time of a heavy particle is

$$\left\langle \delta p^2 \right\rangle = \frac{2n}{m}\left\langle \int q^3 (1 - \cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega \right\rangle , \tag{1.184}$$

where the average has to be taken over the momentum distribution of the light particles.

?

Can you confirm that the differential cross section for light point particles scattered by a hard sphere is given by (1.185)?

Suppose that the heavy particles can be considered as hard spheres with radius R, while the light particles approximate point masses. Then, in the idealised situation of light particles bouncing off heavy, hard spheres,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{R^2}{4} , \quad \int (1 - \cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = \frac{\pi R^2}{2}\int_{-1}^{1}(1 - \cos\theta)\mathrm{d}(\cos\theta) = \pi R^2 , \tag{1.185}$$

and the diffusion coefficient D_2 becomes

$$D_2 = \frac{1}{2}\left\langle \delta p^2 \right\rangle = \frac{\pi n R^2}{m}\left\langle q^3 \right\rangle , \tag{1.186}$$

where the average over the cubed momentum of the light particles remains.

If their velocity distribution is of Maxwellian form with temperature $\bar{T}$,

$$\left\langle q^3 \right\rangle = \frac{8\sqrt{2}}{\sqrt{\pi}}(mk\bar{T})^{3/2} \ , \tag{1.187}$$

and the diffusion coefficient finally assumes the form

$$D_2 = 8nR^2 \left[2\pi m \left(k\bar{T}\right)^3\right]^{1/2} \tag{1.188}$$

which is even independent of the momentum p. This result can now be used with the Fokker-Planck equation (1.180) to calculate how a non-equilibrium phase-space distribution f evolves in time towards its equilibrium by collisions with heavier particles.

Suggested further reading: [1, 2, 3, 4]

Chapter 2

Radiation Processes

This chapter deals with radiation processes. These are defined as processes by which electromagnetic radiation is either scattered, emitted or absorbed by matter. In the first five sections, radiation will be treated as a classical electromagnetic wave. We shall begin with the very illustrative case of Thomson scattering, then give a general description of spectra, proceed to synchrotron radiation and bremsstrahlung and finally consider the drag that a charged particle experiences as it moves through a radiation field. Up to that point, our main theoretical instrument will be Larmor's formula, either in its fully relativistic form (1.138) or in its non-relativistic approximation (1.142), which quantifies the radiation power of a charge moving with a velocity $\vec{\beta}$ and accelerated by $\dot{\vec{\beta}}$. Then, we shall leave the classical picture of electromagnetic waves and consider quantum properties of radiation. The theory of Compton scattering treats electromagnetic radiation as a stream of photons. Emission of radiation by quantum systems will be discussed next, treating their interaction with electromagnetic radiation at a semi-classical, perturbative level, i.e. without quantisation of the electromagnetic field. This will lead us to the calculation of radiative transition probabilities and finally to the shape of spectral lines.

2.1 Thomson scattering

Thomson scattering describes perhaps the simplest case of interaction between an electromagnetic wave and a point charge: The wave accelerates the charge transversally to its propagation direction. Due to its accelerated motion, the charge radiates according to the non-relativistic Larmor formula. The emitted radiation power, divided by the flux density of the incoming radiation, is the Thomson cross section. Its differential, polarisation-dependent or polarisation-averaged forms (2.13) and (2.14) as well as the total Thomson cross section (2.15) are the main results of this section.

Let us begin with a monochromatic, polarised, plane electromagnetic wave hitting an electron at rest. By the Lorentz force, it will accelerate the electron to move harmonically. Because of this accelerated motion, the electron will radiate according to Larmor's formula, as we have seen in Sect. 1.3.5. We ask now how the energy radiated by the electron relates to the energy transported by the incoming wave.

For definiteness, we introduce a coordinate frame such that the infalling electromagnetic wave propagates into the $\hat{e}_z$ direction. The $\vec{E}$ and $\vec{B}$ vectors must then fall into the x-y plane because electromagnetic waves in vacuum are transversal. The polarisation angle will be fixed below. We place the electron at rest into the origin of the coordinate frame.

?

Why are electromagnetic waves in vacuum transversal? Can you construct situations in which longitudinal electromagnetic waves occur?

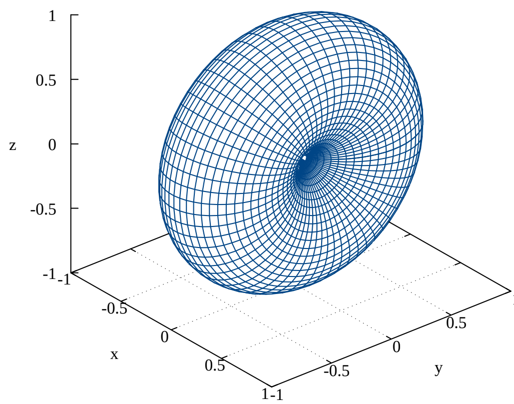

Figure 2.1 The spatial radiation pattern of a non-relativistic charge accelerated along the x axis is shown here. (The x axis points horizontally towards the bottom right).

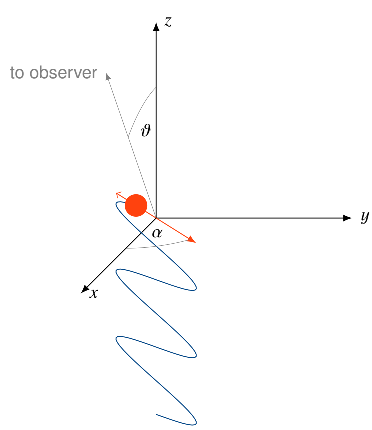

Figure 2.2 Choice of the coordinate system for the treatment of Thomson scattering in the text.

The electron experiences the Lorentz force

$$m_e\ddot{\vec{x}} = -e\vec{E} - \frac{e}{c}\vec{v} \times \vec{B} = -e\left(\vec{E} + \vec{\beta} \times \vec{B}\right) \ . \tag{2.1}$$

For the incoming wave, $|\vec{B}| = |\vec{E}|$. If the electron moves non-relativistically, $|\vec{v}| \ll c$, the magnetic contribution to the Lorentz force can be neglected since the $\vec{E}$ and $\vec{B}$ fields of an electromagnetic wave in vacuum have equal magnitude. The equation of motion for the electron then reduces to

$$\ddot{\vec{x}} = c\dot{\vec{\beta}} = -\frac{e}{m_\mathrm{e}}\vec{E} \ . \tag{2.2}$$

The non-relativistic limit of the Larmor formula (1.138) is

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{e^2}{4\pi c}\left|\hat{e}\times\dot{\vec{\beta}}\right|^2 \ . \tag{2.3}$$

It gives the energy radiated per unit time into the solid-angle element $\mathrm{d}\Omega$ around the vector $\hat{e}$ pointing from the charge to the observer (Figure 2.1). Since the electron's motion is non-relativistic, retardation effects can be neglected. Inserting the acceleration by the Lorentz force (2.2) with (2.3) gives

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{e^4}{4\pi m_\mathrm{e}^2 c^3}\left|\hat{e}\times\vec{E}\right|^2 \ . \tag{2.4}$$

We rotate the coordinate frame (Figure 2.2) such that the observer lies in the x-z plane,

$$\hat{e} = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix} , \tag{2.5}$$

and introduce the polarisation angle α of the incoming $\vec{E}$ field as the angle enclosed by the $\vec{E}$ vector with the $\hat{e}_x$ axis,

$$\vec{E} = E\begin{pmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{pmatrix} . \tag{2.6}$$

With this choice, we find

$$\hat{e}\times\vec{E} = E\begin{pmatrix} -\sin\alpha\cos\theta \\ \cos\alpha\cos\theta \\ \sin\alpha\sin\theta \end{pmatrix} , \tag{2.7}$$

and the radiated power per solid angle given by (2.4) turns into

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{e^4 E^2}{4\pi m_\mathrm{e}^2 c^3}\left(1-\sin^2\theta\cos^2\alpha\right) \ . \tag{2.8}$$

The infalling energy current density is quantified by the Poynting vector of the incoming wave,

$$\vec{S} = \frac{c}{4\pi}\left|\vec{E}\right|^2\hat{e}_z \ . \tag{2.9}$$

This is the energy per unit area and unit time impinging on the electron. The ratio between the energy radiated per unit time and unit solid angle and the energy current density,

$$\frac{1}{|\vec{S}|}\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{e^4}{m_\mathrm{e}^2 c^4}\left(1-\sin^2\theta\cos^2\alpha\right) \ , \tag{2.10}$$

has the dimension of an area. It is the *differential Thomson cross section* for polarised light (Figure 2.3a).

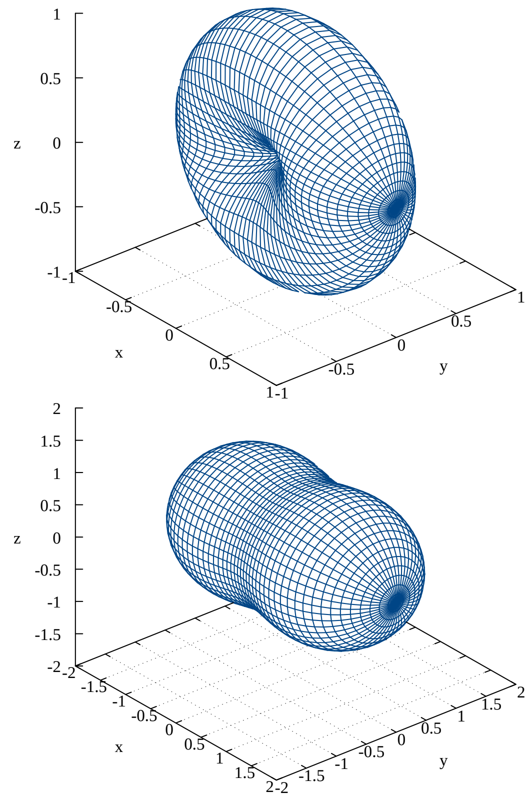

Figure 2.3 These bodies illustrate the polarised and the unpolarised Thomson cross sections for electromagnetic waves propagating along the positive x direction. *Top panel*: The directional dependence of the polarised Thomson cross section on the scattering angle is shown here, with the polarisation angle being the angle enclosed with the z axis (i.e. the polar angle). *Bottom panel*: The unpolarised Thomson cross section is forward-backward symmetric.

The prefactor $e^4/m_e^2c^4$ has an interesting and intuitive meaning. Suppose we want to explain the entire rest-energy of the electron by the electrostatic energy of the charge e distributed over a sphere of radius r_e. We would then require

$$m_e c^2 = \frac{e^2}{r_e} \tag{2.11}$$

and find the classical electron radius

$$r_e = \frac{e^2}{m_e c^2} \approx 2.81 \cdot 10^{-13}\,\mathrm{cm}\ . \tag{2.12}$$

For ions composed of N nucleons and having a charge number Z, this classical radius is at least approximately $Z^2/(1800\,N)$ times smaller because of their much higher mass. The Thomson cross section of ions is therefore generally negligibly small compared to that of the electrons. Electromagnetic radiation flowing through, say, a hydrogen plasma is scattered by the electrons, which then interact mainly by Coulomb collisions with the ions.

The classical electron radius brings the differential, polarised Thomson cross section (2.10) into the simple, intuitive form

$$\frac{d\sigma}{d\Omega} = r_e^2 \left(1 - \sin^2\theta \cos^2\alpha\right) \ . \tag{2.13}$$

?

Before moving on, verify (2.13), then average over polarisation angles and integrate over the solid angle.

For unpolarised light, we need to average (2.13) over all polarisation angles α. This average leads to the unpolarised, differential Thomson cross section (Figure 2.3b)

$$\left\langle\frac{d\sigma}{d\Omega}\right\rangle_\alpha = \frac{1}{2\pi}\int_0^{2\pi}\frac{d\sigma}{d\Omega}d\alpha = \frac{r_e^2}{2}\left(2 - \sin^2\theta\right) = \frac{r_e^2}{2}\left(1 + \cos^2\theta\right) \ . \tag{2.14}$$

If we finally integrate over all directions into which the radiation is scattered, we find the *total Thomson cross section*

$$\begin{aligned}\sigma_T &= \int d\Omega \left\langle\frac{d\sigma}{d\Omega}\right\rangle_\alpha = \pi r_e^2 \int_{-1}^{1} d(\cos\theta)\left(1 + \cos^2\theta\right) \\ &= \frac{8\pi}{3} r_e^2 \approx 6.64 \cdot 10^{-25}\,\text{cm}^2 \ . \end{aligned} \tag{2.15}$$

This can be interpreted as the area that a single, non-relativistic electron puts in the way of incoming, unpolarised radiation.

Problems

1. Work out the mean molecular mass for a mixture of neutral, atomic hydrogen and helium as a function of the hydrogen mass fraction X.
2. Consider an electron at the origin of the coordinate system, illuminated by two unpolarised electromagnetic wave bundles propagating along the $-y$ and $-z$ axes with different energy current densities S_y and S_z.
   (a) Find the radiation power radiated into the x direction.
   (b) Is the scattered radiation polarised?

2.2 Spectra

This brief section discusses how electromagnetic spectra of accelerated charges can be computed. The starting point is Larmor's equation in its relativistic or non-relativistic forms, giving the radiation power. The total energy radiated away is the time integral over the power which, by Parseval's equation for Fourier-conjugate functions, can be converted to a frequency integral. Its integrand is the energy per unit frequency, i.e. the spectrum. This allows us to derive the fully relativistic expression (2.36) for the spectrum. The substantially simplified versions (2.39) and (2.42) for non-relativistic charges can be directly derived from the non-relativistic Larmor equation.

Example: Eddington Luminosity

Let us immediately apply the Thomson scattering cross section to the following situation. Suppose ionised gas surrounds a hot, spherically-symmetric, radiating body of mass M. The radiation carries the momentum current density

$$\frac{\vec{S}}{c} = \frac{1}{4\pi} \left|\vec{E}^2\right| \hat{e} \; , \tag{2.16}$$

given by the components of Maxwell's stress-energy tensor. Since this is the momentum flowing per unit time through unit area, it corresponds to a force per unit area, or a pressure exerted on an ideally absorbing wall.

The total energy emitted by the star per unit time is its luminosity L. Exploiting the spherical symmetry, we have

$$L = \int \vec{S} \cdot \mathrm{d}\vec{a} = 4\pi R^2 \cdot \frac{c}{4\pi} \left|\vec{E}(R)\right|^2 \; , \tag{2.17}$$

where $\vec{E}(R)$ is the electric field strength at radius R. According to (2.16), the radiation pressure there is expressed by L after eliminating the electric field $\vec{E}$,

$$\frac{\vec{S}}{c} = \frac{L}{4\pi c R^2} \hat{e} \; . \tag{2.18}$$

Each electron in the surrounding plasma has a Thomson-scattering cross section of σ_T and thus experiences the force

$$\vec{F}_R = \frac{\vec{S}}{c} \cdot \sigma_\mathrm{T} = \frac{L}{4\pi R^2 c} \sigma_\mathrm{T} \hat{e} \tag{2.19}$$

by the radiation pressure. Recall that the force on the ions in the plasma is lower by a factor of $\approx Z^2/(1800\,N)$ if the ions have the charge Ze and are composed of N nucleons. This radiation-pressure force acting radially outward is counter-acted by the gravitational force of the mass M of the central body,

$$\vec{F}_G = -\frac{GMm}{R^2} \hat{e} \; . \tag{2.20}$$

Both forces compensate each other if the luminosity L satisfies

$$\frac{L}{4\pi R^2 c} \sigma_\mathrm{T} = \frac{GMm}{R^2} \; , \tag{2.21}$$

i.e. if the luminosity reaches the *Eddington limit*

$$L = L_\mathrm{Edd} = \frac{4\pi GMm}{\sigma_\mathrm{T}} c \; . \tag{2.22}$$

Inserting a solar mass for M and a proton mass for m here results in

$$L_\mathrm{Edd} = 1.26 \cdot 10^{38} \, \frac{\mathrm{erg}}{\mathrm{s}} = 3.28 \cdot 10^4 \, L_\odot \tag{2.23}$$

(see Tabs. 1.3 and 1.4). ◂

Example: Eddington Luminosity (continued)

Note that we have deliberately not specified the particle mass m in (2.20) and the following equations to be the electron mass. Consider a hydrogen plasma consisting of an equal mixture of electrons and protons. By (2.12), the Thomson cross section of a proton is about $1800^2 \approx 3.2 \cdot 10^6$ times smaller than that of an electron. However, while essentially only the electrons feel the radiation pressure, they are tighly coupled by Coulomb interactions to the protons. The radiation-pressure force thus needs to compensate the gravitational force felt by the electrons and the protons together. The particle mass m inserted in (2.22) should therefore be the total mass per electron rather than the electron mass alone. For a fully ionised hydrogen plasma, we can approximate m by the proton mass m_p. ◂

The energy received by an observer from a radiating electron, flowing into the solid angle $\mathrm{d}\Omega$, is

$$\frac{\mathrm{d}E}{\mathrm{d}\Omega} = \int \mathrm{d}t\, \frac{\mathrm{d}P}{\mathrm{d}\Omega}\,, \tag{2.24}$$

where $\mathrm{d}P/\mathrm{d}\Omega$ is given by the Larmor formula (1.138). Often, we are interested in the radiation spectrum, i.e. in the distribution of the energy over frequency rather than time. Realising that the time t and the frequency ω are Fourier conjugates, this is most easily found using Plancherel's theorem,

$$\int_{-\infty}^{\infty} \mathrm{d}t\, |f(t)|^2 = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi}\, \left|\hat f(\omega)\right|^2\,, \tag{2.25}$$

which specialises Parseval's equation for continuous Fourier transforms. It relates the integral over a function to that over its Fourier transform. The negative frequencies ω in (2.25) may appear strange here. Nonetheless, they obtain a well-defined meaning because we require that $f(t)$ be real. Then, its Fourier transform $\hat f(\omega)$ must satisfy the relation $\hat f(-\omega) = \hat f^*(\omega)$.

Applying Plancherel's theorem to (2.24) and inserting the Larmor formula (1.138), we find

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}\Omega} &= \frac{e^2}{4\pi c}\int_{-\infty}^{\infty}\mathrm{d}t\left|\frac{\hat e\times\left[\left(\hat e-\vec\beta\right)\times\dot{\vec\beta}\right]}{\left(1-\hat e\cdot\vec\beta\right)^3}\right|^2 = \frac{e^2}{4\pi c}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\left|\hat f(\omega)\right|^2 \\ &= \int_{-\infty}^{\infty}\mathrm{d}\omega\,\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega}\,,\end{aligned} \tag{2.26}$$

where $\hat f(\omega)$ is now specified to be the Fourier transform of the function

$$f(t) = \frac{\hat e\times\left[\left(\hat e-\vec\beta\right)\times\dot{\vec\beta}\right]}{\left(1-\hat e\cdot\vec\beta\right)^3} \tag{2.27}$$

that can directly be read off the Larmor formula (1.138). The spectrum is then given by the absolute square of $\hat f(\omega)$,

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} = \frac{e^2}{8\pi^2 c}\left|\hat f(\omega)\right|^2\,. \tag{2.28}$$

The Fourier transform

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t \, \frac{\hat{e} \times \left[\left(\hat{e} - \vec{\beta}\right) \times \dot{\vec{\beta}}\right]}{\left(1 - \hat{e} \cdot \vec{\beta}\right)^3} \, \mathrm{e}^{-\mathrm{i}\omega t} \tag{2.29}$$

simplifies considerably realising that the integrand needs to be evaluated at the retarded time $t' = t - R/c$, where R is the distance from the observer to the electron at the retarded time. Taking into account that the differential $\mathrm{d}t'$ of the retarded time is related to $\mathrm{d}t$ by (1.136), we can first cancel one factor $(1 - \hat{e} \cdot \vec{\beta})$ from the denominator and write

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t' \, \frac{\hat{e} \times \left[\left(\hat{e} - \vec{\beta}\right) \times \dot{\vec{\beta}}\right]}{\left(1 - \hat{e} \cdot \vec{\beta}\right)^2} \, \mathrm{e}^{-\mathrm{i}\omega(t'+R/c)} \,. \tag{2.30}$$

Furthermore, a short calculation shows that the integrand can be written as a total derivative with respect to the retarded time t',

$$\frac{\hat{e} \times \left[\left(\hat{e} - \vec{\beta}\right) \times \dot{\vec{\beta}}\right]}{\left(1 - \hat{e} \cdot \vec{\beta}\right)^2} = \frac{\mathrm{d}}{\mathrm{d}t'} \left[\frac{\hat{e} \times \left(\hat{e} \times \vec{\beta}\right)}{\left(1 - \hat{e} \cdot \vec{\beta}\right)}\right] \,. \tag{2.31}$$

?

Verify (2.31) by your own calculation.

This leaves us with

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t' \, \frac{\mathrm{d}}{\mathrm{d}t'} \left[\frac{\hat{e} \times \left(\hat{e} \times \vec{\beta}\right)}{\left(1 - \hat{e} \cdot \vec{\beta}\right)}\right] \mathrm{e}^{-\mathrm{i}\omega(t'+R/c)} \,, \tag{2.32}$$

which calls for partial integration. Before we get to that, however, we decompose the distance vector $\vec{R}$ from the radiating electron to the observer into the distance vector $\vec{x}$ from the center of the orbit to the electron and the distance vector $\vec{r}$ from the center of the orbit to the observer,

$$\vec{R} = \vec{r} - \vec{x} \,. \tag{2.33}$$

The idea behind this decomposition is that the motion of the radiating charge is confined to a distant source, and thus to a volume which is far away and small compared to its distance from the observer. The retarded distance R is then $\hat{e} \cdot \vec{R}$, and its derivative with respect to the retarded time is

$$\frac{\mathrm{d}}{\mathrm{d}t'} \frac{R}{c} = -\hat{e} \cdot \vec{\beta} \,; \tag{2.34}$$

compare (1.126). Assuming that the emission in the distant past and in the far future can be neglected, we can ignore the boundary terms appearing in the partial integration of (2.32). Taking (2.34) into account, the partial integration gives

$$\begin{aligned} \hat{f}(\omega) &= -\int_{-\infty}^{\infty} \mathrm{d}t' \left[\frac{\hat{e} \times \left(\hat{e} \times \vec{\beta}\right)}{\left(1 - \hat{e} \cdot \vec{\beta}\right)}\right] \frac{\mathrm{d}}{\mathrm{d}t'} \, \mathrm{e}^{-\mathrm{i}\omega(t'+R/c)} \\ &= \mathrm{i}\omega \int_{-\infty}^{\infty} \mathrm{d}t' \left[\hat{e} \times \left(\hat{e} \times \vec{\beta}\right)\right] \mathrm{e}^{-\mathrm{i}\omega(t'-\hat{e}\cdot\vec{x}/c)} \,, \end{aligned} \tag{2.35}$$

where we have ignored the constant phase factor $\mathrm{e}^{\mathrm{i}\omega\hat{e}\cdot\vec{r}}$. It is irrelevant because we later need to take the absolute value of $\hat{f}(\omega)$ anyway.

It is worth noting again that we have made a single approximation in the preceding calculation which is perfectly legitimate in typical astrophysical situations: We have assumed that the radiating electron is confined to a distant volume that is small compared to its distance from the observer. This has allowed us to derive a general prescription for calculating radiation spectra, expressed by (2.29) with $\hat{f}(\omega)$ given by the Fourier transform (2.35),

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} = \frac{e^2}{8\pi^2c}\left|\hat{f}(\omega)\right|^2 = \frac{e^2\omega^2}{8\pi^2c}\left|\int_{-\infty}^{\infty}\mathrm{d}t'\,\left[\hat{e}\times\left(\hat{e}\times\vec{\beta}\right)\right]\,\mathrm{e}^{-\mathrm{i}\omega(t'-\hat{e}\cdot\vec{x}/c)}\right|^2 , \tag{2.36}$$

understanding that the integrand has to be evaluated at the retarded time t' and that $\vec{x} = \vec{x}(t')$ describes the electron's orbit about a fixed reference point within the volume it is confined to. We can now apply this general result to different circumstances relevant in astrophysics.

The calculation simplifies considerably for non-relativistically moving charges. Then, relativistic beaming is irrelevant, retardation effects can be ignored, and terms of higher than linear order in β and $\dot{\beta}$ can be neglected. We can then begin with the direction-integrated, non-relativistic Larmor formula following from (1.141) by setting $\gamma = 1$ and dropping the fourth-order term in β. Then,

$$E = \int_{-\infty}^{\infty} P\,\mathrm{d}t = \frac{2e^2}{3c^3}\int_{-\infty}^{\infty}\left|\vec{a}(t)\right|^2\mathrm{d}t , \tag{2.37}$$

where $\vec{a} = \ddot{\vec{x}}$ is the acceleration experienced by the charge. Employing Plancherel's theorem (2.25) once more, we can continue writing (2.37) as

$$E = \frac{2e^2}{3c^3}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\left|\hat{\vec{a}}(\omega)\right|^2 = \int_{-\infty}^{\infty}\mathrm{d}\omega\,\frac{\mathrm{d}E}{\mathrm{d}\omega} , \tag{2.38}$$

which yields the non-relativistic, direction-integrated spectrum

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{e^2}{3\pi c^3}\left|\hat{\vec{a}}(\omega)\right|^2 . \tag{2.39}$$

The Fourier transform $\hat{\vec{a}}$ of the acceleration can easily be expressed by the Fourier transform of the orbit itself. Since

$$\vec{a} = \ddot{\vec{x}} = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\,\hat{\vec{x}}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t} = -\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\,\omega^2\hat{\vec{x}}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t} , \tag{2.40}$$

the Fourier transform of $\hat{\vec{a}}$ is

$$\hat{\vec{a}} = -\omega^2\hat{\vec{x}}(\omega) , \tag{2.41}$$

which allows us to calculate the spectrum directly from

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{e^2\omega^4}{3\pi c^3}\left|\hat{\vec{x}}(\omega)\right|^2 . \tag{2.42}$$

Example: Electron on a circular orbit

A quite simple example is an electron orbiting on a circle of radius r with an angular frequency ω_0. Since its orbit is given by

$$\vec{x}(t) = r \begin{pmatrix} \cos\omega_0 t \\ \sin\omega_0 t \\ 0 \end{pmatrix} , \quad x_1(t) + \mathrm{i}x_2(t) = r\mathrm{e}^{\mathrm{i}\omega_0 t} , \tag{2.43}$$

the Fourier transform of x_1 and x_2 together is

$$\hat{x}_1(\omega) + \mathrm{i}\hat{x}_2(\omega) = r\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(\omega_0-\omega)t}\,\mathrm{d}t = 2\pi r\,\delta_\mathrm{D}(\omega_0-\omega) . \tag{2.44}$$

Its spectrum is thus a single, sharp line emitting the energy

$$\int_{-\infty}^{\infty}\mathrm{d}\omega\,\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{4\pi}{3}\frac{e^2r^2\omega_0^4}{c^3} . \tag{2.45}$$

◀

Example: Electron under constant acceleration

For another illustrative example, suppose an electron is accelerated with constant acceleration $\vec{a}$ during a finite time interval $-\tau/2 \le t \le \tau/2$. The Fourier transform of this acceleration is

$$\hat{\vec{a}}(\omega) = \vec{a}\int_{-\tau/2}^{\tau/2}\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}t = -\frac{\mathrm{i}\vec{a}}{\omega}\left(\mathrm{e}^{\mathrm{i}\omega\tau/2} - \mathrm{e}^{-\mathrm{i}\omega\tau/2}\right) = \frac{2\vec{a}}{\omega}\sin\frac{\omega\tau}{2} , \tag{2.46}$$

which we can insert directly into (2.39) to find the spectrum

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{4e^2\vec{a}^2}{3\pi c^3\omega^2}\sin^2\frac{\omega\tau}{2} . \tag{2.47}$$

◀

Problems

1. Verify equation (2.31).

2. Let the Fourier transform of a function $f(x)$ and the inverse transform of its Fourier conjugate $\hat{f}(k)$ be defined by

$$\hat{f}(k) = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\mathrm{e}^{\mathrm{i}kx} \,, \quad f(x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\, \hat{f}(k)\mathrm{e}^{-\mathrm{i}kx} \,. \tag{2.48}$$

Prove the following identities:

(a)

$$\hat{f}^*(-k) = \hat{f}(k) \tag{2.49}$$

for real functions, $f(x) \in \mathbb{R}$.

(b)

$$\widehat{f * g} = \hat{f}\,\hat{g} \tag{2.50}$$

if

$$(f * g)(x) := \int_{-\infty}^{\infty} \mathrm{d}y f(x - y)g(y) \tag{2.51}$$

is the convolution of the two functions f and g.

(c)

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \hat{f}(k)\hat{g}^*(k) = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)g^*(x) \tag{2.52}$$

(Parseval's equation).

(d)

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} |\hat{f}(k)|^2 = \int_{-\infty}^{\infty} \mathrm{d}x\, |f(x)|^2 \,. \tag{2.53}$$

3. Consider an electron whose one-dimensional trajectory $x(t)$ satisfies the differential equation of a damped harmonic oscillator,

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0 \,. \tag{2.54}$$

(a) What is the oscillator frequency ω if ω_0 is the system's eigenfrequency? *Hint*: Try the ansatz $x(t) \propto \mathrm{e}^{\pm\mathrm{i}\omega t}$. What does a complex frequency mean physically?

(b) Show that the solution of the differential equation is given by

$$x(t) = \frac{v_0}{\bar{\omega}}\mathrm{e}^{-\gamma t} \sin\bar{\omega}t \quad \text{with} \quad \bar{\omega} = \sqrt{\omega_0^2 - \gamma^2} \,. \tag{2.55}$$

if $\omega_0 > \gamma$ and the initial conditions are $x(t = 0) = 0$ and $\dot{x}(t = 0) = v_0$.

(c) Calculate the Fourier transform $\hat{x}(\omega)$. Assume that $x(t) = 0$ for $t < 0$.

(d) Calculate the spectrum $\mathrm{d}E/\mathrm{d}\omega$ of the moving electron.

(e) What does the spectrum look like if both $\omega \gg \omega_0$ and $\omega \gg \gamma$?

2.3 Synchrotron radiation

In this section, the power and the spectrum radiated by a relativistic charge gyrating in a magnetic field are calculated. This is entirely an application of Larmor's equation from classical electrodynamics and the general formulae for calculating spectra derived in the preceding section. We shall first consider the trajectory of the electron, then discuss relativistic beaming and its effects, and proceed directly to the synchrotron power in (2.68) and the synchrotron spectrum in (2.86). The main assumptions are that the emitting charge is confined to a volume whose dimensions are small compared to its distance from the observer and that the source is ultra-relativistic. Besides the shape of the synchrotron spectrum, an important result of the discussion is that relativistic beaming allows the observer to see the signal only during a very short time per orbit, which substantially broadens the spectrum since frequency and time are Fourier conjugates.

2.3.1 Larmor frequency and relativistic focussing

Consider now an electron moving relativistically in a homogeneous magnetic field. Without electric field, $\vec{E} = 0$, the Lorentz force (1.146) causes the acceleration

$$\frac{\mathrm{d}\,(\gamma\vec{v})}{\mathrm{d}t} = -\frac{e}{mc}\vec{v}\times\vec{B}\;. \tag{2.56}$$

Since this purely magnetic Lorentz force is perpendicular to the velocity, it cannot change the electron's energy, thus $\gamma = \text{const}$. Let us rotate the coordinate frame such that $\vec{B}$ is aligned with the z axis, hence $\vec{B} = B\hat{e}_z$. Then,

$$\frac{\mathrm{d}(\gamma v_z)}{\mathrm{d}t} = \gamma\dot{v}_z = 0 \tag{2.57}$$

while

$$\dot{v}_x = -\frac{eB}{\gamma mc}v_y\;, \quad \dot{v}_y = \frac{eB}{\gamma mc}v_x\;. \tag{2.58}$$

Taking a second time derivative of either of the two equations (2.58) and combining it with the respective other equation gives

$$\ddot{v}_i + \left(\frac{eB}{\gamma mc}\right)^2 v_i = 0\;, \quad i = x, y\;. \tag{2.59}$$

This is the equation of a harmonic oscillator with the Larmor frequency

$$\omega_\mathrm{L} = \frac{eB}{\gamma mc} = 17.6\,\mathrm{Hz}\,\gamma^{-1}\left(\frac{B}{\mu\mathrm{G}}\right)\left(\frac{m_\mathrm{e}}{m}\right)\;. \tag{2.60}$$

In a constant magnetic field, the electron therefore describes a circular orbit with cyclic frequency ω_L in the plane perpendicular to $\vec{B}$, while it moves with constant velocity along $\vec{B}$ (Figure 2.4). If it has $v_z \neq 0$ initially, it orbits on a helix with constant radius and pitch angle.

Let us now assume for simplicity that $v_z = 0$ so that the electron moves on a circle in the plane perpendicular to $\vec{B}$ (Figure 2.5). Alternatively, we can transform into a reference frame co-moving with the mean motion of the

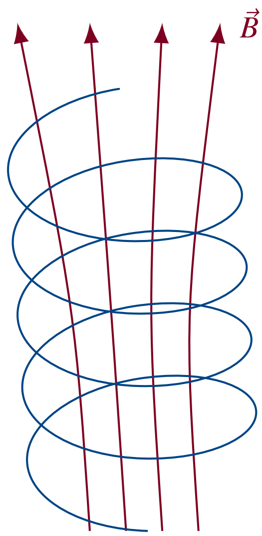

Figure 2.4 The trajectory of a charge in a locally constant magnetic field $\vec{B}$ is a helix.

electron. On a circular orbit, the acceleration is perpendicular to the velocity, $\vec{\beta} \perp \dot{\vec{\beta}}$ or $\vec{\beta} \cdot \dot{\vec{\beta}} = 0$. Since the electron is supposed to move relativistically, $\beta \approx 1$, and we can approximate

$$1 - \beta = \frac{1 - \beta^2}{1 + \beta} \approx \frac{1}{2\gamma^2}\,, \quad \beta \approx 1 - \frac{1}{2\gamma^2}\,. \tag{2.61}$$

?

Why is it appropriate and consistent to approximate $1 + \beta \approx$ in (2.61)?

Introducing the angle θ between $\hat{e}$ and $\vec{\beta}$ by $\beta \cos\theta = \hat{e} \cdot \vec{\beta}$, we see that the factor $(1 - \hat{e} \cdot \vec{\beta})^{-1} = (1 - \beta\cos\theta)^{-1}$ in the Larmor formula (1.138) is very large. In the direction of the motion, $\theta = 0$,

$$\left(1 - \hat{e} \cdot \vec{\beta}\right)^{-1} = (1 - \beta)^{-1} \approx 2\gamma^2\,, \tag{2.62}$$

and the factor $(1 - \beta\cos\theta)^{-1}$ drops to half its maximum within a narrow angle. Requiring

$$\frac{1}{1 - \hat{e} \cdot \vec{\beta}} \gtrsim \frac{1}{2(1 - \beta)}\,, \tag{2.63}$$

we find the condition

$$\cos\theta \approx 1 - \frac{\theta^2}{2} \gtrsim 2 - \frac{1}{\beta}\,, \tag{2.64}$$

from which we can read off

$$\theta \lesssim \sqrt{2\left(\frac{1}{\beta} - 1\right)} \approx \sqrt{2(1 - \beta)} \approx \frac{1}{\gamma}\,. \tag{2.65}$$

The energy radiated by the electron is thus confined to a very narrow beam with opening angle $\lesssim \gamma^{-1}$. This will allow us to introduce several well-justified approximations as we go along.

2.3.2 Synchrotron power

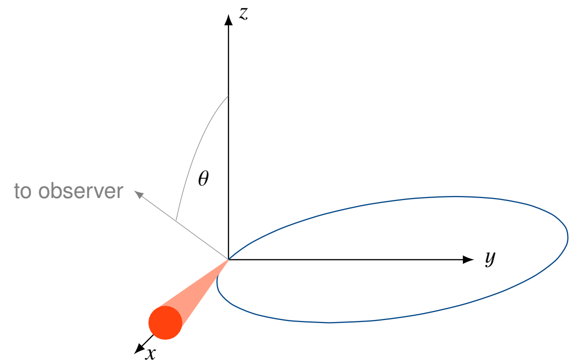

Figure 2.5 Illustration of how the coordinate frame is chosen for the calculation of the synchrotron power and the synchrotron spectrum carried out in the text.

Let us first introduce a coordinate frame oriented such that the electron's orbit falls into the x-y plane, while the observer is in the x-z plane. Furthermore, we shift the coordinate origin into the centre of the circular orbit and choose the zero point of the retarded time t' such that the electron moves into the $\hat{e}_x$ direction at $t' = 0$. Then, we can write

$$\hat{e} = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix} , \quad \vec{x} = x \begin{pmatrix} \sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} ,$$
$$\vec{\beta} = \beta \begin{pmatrix} \cos\varphi \\ -\sin\varphi \\ 0 \end{pmatrix} , \quad \dot{\vec{\beta}} = \beta\dot{\varphi} \begin{pmatrix} -\sin\varphi \\ -\cos\varphi \\ 0 \end{pmatrix} , \tag{2.66}$$

where x is the radius of the orbit and the dimension-less velocity is $\beta = x\dot{\varphi}/c = x\omega_L/c$.

The total synchrotron power follows directly from the integrated Larmor formula (1.141). Since $\vec{\beta} \perp \dot{\vec{\beta}}$ in the case of synchrotron radiation, we first obtain

$$P = \frac{2e^2}{3c}\gamma^6 \left[\dot{\beta}^2 - \left(\vec{\beta} \times \dot{\vec{\beta}}\right)^2\right] = \frac{2e^2}{3c}\gamma^6\dot{\beta}^2\left(1 - \beta^2\right) = \frac{2e^2}{3c}\gamma^4\dot{\beta}^2 \ . \tag{2.67}$$

Since $\dot{\beta} = \beta\dot{\varphi} \approx \omega_L$, we can further simplify

$$P = \frac{2e^2}{3c}\gamma^4\omega_L^2 = \frac{2e^2}{3c}\gamma^4\left(\frac{eB}{\gamma mc}\right)^2 = \frac{8\pi}{3}r_e^2 c\gamma^2\frac{B^2}{4\pi} = c\gamma^2\sigma_T U_B \ , \tag{2.68}$$

where $U_B = B^2/4\pi$ is the energy density in the magnetic field that can be read off Maxwell's energy-momentum tensor, see (1.112), and σ_T is the Thomson cross section, derived in the non-relativistic regime. As we shall see later, this is a very intuitive expression for the total synchrotron power.

2.3.3 Synchrotron spectrum

We now turn to the evaluation of the spectrum (2.36) under the given circumstances. Expanding first the double vector product in (2.36), we find

$$\hat{e} \times \left(\hat{e} \times \vec{\beta}\right) = \left(\hat{e} \cdot \vec{\beta}\right)\hat{e} - \vec{\beta} = \beta \begin{pmatrix} -\cos\varphi\cos^2\theta \\ \sin\varphi \\ \cos\varphi\sin\theta\cos\theta \end{pmatrix} . \tag{2.69}$$

This vector must be perpendicular to the line-of-sight, whose direction is given by $\hat{e}$. We can thus expand it into two basis vectors perpendicular to $\hat{e}$, which we choose to be $\hat{e}_y$ and

$$\hat{e}_\perp = \hat{e} \times \hat{e}_y = \begin{pmatrix} -\cos\theta \\ 0 \\ \sin\theta \end{pmatrix} . \tag{2.70}$$

In this basis,

$$\hat{e} \times \left(\hat{e} \times \vec{\beta}\right) = \beta\cos\varphi\cos\theta\hat{e}_\perp + \beta\sin\varphi\hat{e}_y . \tag{2.71}$$

The phase ψ of the exponential in (2.36) is

$$\omega\left(t' - \frac{\hat{e}\cdot\vec{x}}{c}\right) =: \psi = \omega\left(t' - \frac{x\sin\theta\sin\varphi}{c}\right) . \tag{2.72}$$

We can now make use of the fact that the radiation of the relativistically moving electron is strongly focussed into its forward direction, since the opening angle of the radiation cone is approximately confined to $[-\gamma^{-1}, \gamma^{-1}]$, as we have discussed before. This implies that our observer will see the radiation only when $|\varphi| \lesssim \gamma^{-1}$ and $|\theta - \pi/2| \lesssim \gamma^{-1}$. Since $\gamma \gg 1$, the angle θ is close to $\pi/2$. We introduce its complement $\theta \equiv \pi/2 - \theta \ll 1$ and approximate

$$\sin\theta = \sin\left(\frac{\pi}{2} - \theta\right) = \cos\theta \approx 1 + \frac{\theta^2}{2} , \quad \cos\theta = \cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta \approx \theta . \tag{2.73}$$

The expansion in φ is effectively an expansion in t', for $\varphi = \omega_\mathrm{L} t'$. We shall see later that we need to carry it to order t'^3, hence

$$\sin\varphi \approx \omega_\mathrm{L} t'\left(1 - \frac{\omega_\mathrm{L}^2 t'^2}{6}\right) , \quad \cos\varphi \approx 1 - \frac{\omega_\mathrm{L}^2 t'^2}{2} . \tag{2.74}$$

We thus have

$$\hat{e} \times \left(\hat{e} \times \vec{\beta}\right) \approx \beta\theta\hat{e}_\perp + \beta\omega_\mathrm{L} t'\hat{e}_y \approx \theta\hat{e}_\perp + \omega_\mathrm{L} t'\hat{e}_y , \tag{2.75}$$

and the Fourier phase becomes

$$\psi \approx \omega t'\left[1 - \beta\left(1 - \frac{\omega_\mathrm{L}^2 t'^2}{6}\right)\left(1 - \frac{\theta^2}{2}\right)\right] \approx \frac{\omega t'}{2}\left(\frac{1}{\gamma^2} + \theta^2 + \frac{\omega_\mathrm{L}^2 t'^2}{3}\right) \tag{2.76}$$

where we have used the relations

$$\beta = \frac{x\omega_\mathrm{L}}{c} , \quad 1 - \beta = \frac{1}{2\gamma^2} , \quad \beta \approx 1 . \tag{2.77}$$

To further simplify the expression for the Fourier phase, we pull the factor $(\gamma^{-2} + \theta^2)$ out of the parenthesis in (2.76) to find first

$$\psi \approx \frac{\omega t'}{2}\left(\frac{1}{\gamma^2} + \theta^2\right)\left(1 + \frac{\tau^2}{3}\right) \tag{2.78}$$

with the new dimension-less time variable

$$\tau := \frac{\omega_\mathrm{L} t'}{\sqrt{\gamma^{-2} + \theta^2}} \ . \tag{2.79}$$

Defining further the dimension-less frequency

$$\xi := \frac{\omega}{3\omega_\mathrm{L}}\left(\frac{1}{\gamma^2} + \theta^2\right)^{3/2} \ , \tag{2.80}$$

we can write the phase as

$$\psi = \frac{3\xi\tau}{2}\left(1 + \frac{\tau^2}{3}\right) \ . \tag{2.81}$$

?

Verify the approximate expression (2.81) for the phase function ψ.

Combining expression (2.75) for the double vector product, inserting the Fourier phase ψ from (2.78) and transforming the integration variable from t' to τ as defined in (2.79), we find that we can split the function $\hat{\vec{f}}(\omega)$ introduced in (2.35) as

$$\hat{\vec{f}}(\omega) = \hat{f}_\perp(\omega)\hat{e}_\perp + \hat{f}_\parallel(\omega)\hat{e}_y \ , \tag{2.82}$$

where the perpendicular and parallel Fourier amplitudes are

$$\begin{aligned} \hat{f}_\perp(\omega) &= -\mathrm{i}\frac{\omega}{\omega_\mathrm{L}}\theta\left(\frac{1}{\gamma^2} + \theta^2\right)^{1/2}\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{e}^{-\mathrm{i}\psi} \ , \\ \hat{f}_\parallel(\omega) &= -\mathrm{i}\frac{\omega}{\omega_\mathrm{L}}\left(\frac{1}{\gamma^2} + \theta^2\right)\int_{-\infty}^{\infty}\tau\mathrm{d}\tau\,\mathrm{e}^{-\mathrm{i}\psi} \ . \end{aligned} \tag{2.83}$$

The remaining integrals are Bessel functions of fractional order,

$$\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{e}^{-\mathrm{i}\psi} = \frac{2}{\sqrt{3}}K_{1/3}(\xi) \ , \quad \int_{-\infty}^{\infty}\tau\mathrm{d}\tau\,\mathrm{e}^{-\mathrm{i}\psi} = -\frac{2\mathrm{i}}{\sqrt{3}}K_{2/3}(\xi) \ . \tag{2.84}$$

Putting these results together, we can express (2.36) as

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} = \frac{e^2}{4\pi c}\left[\left|\hat{f}_\perp(\omega)\right|^2 + \left|\hat{f}_\parallel(\omega)\right|^2\right] \ , \tag{2.85}$$

or, introducing the preceding results for the functions $f_\perp$ and $f_\parallel$,

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} = \frac{e^2\omega^2}{3\pi c\omega_\mathrm{L}^2}\left(\frac{1}{\gamma^2} + \theta^2\right)^2\left[\frac{\theta^2}{\gamma^{-2} + \theta^2}K_{1/3}^2(\xi) + K_{2/3}^2(\xi)\right] \ . \tag{2.86}$$

This is the synchrotron spectrum (Figure 2.6).

To obtain further insight into the shape of the spectrum, let us shift the observer into the orbital plane of the electron. Since the radiation is focussed into a narrow cone with $|\theta| \lesssim \gamma^{-1}$, this is not a strong simplification. We first realise

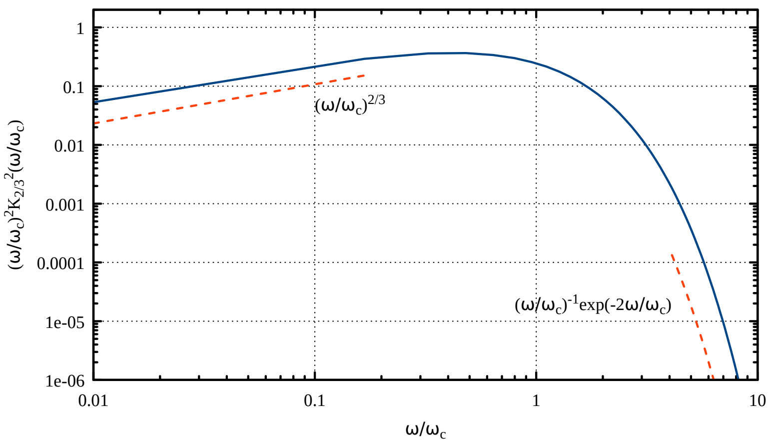

Figure 2.6 The shape of the synchrotron spectrum is shown in arbitrary units as a function of the scaled frequency $\xi = \omega/\omega_c$. The polar angle was set to $\theta = \pi/2$, i.e. this is the spectral shape in the orbital plane of the electron. The Lorentz factor is irrelevant here since it only affects the amplitude, not the shape of the spectrum.

that the intensity of the radiation component polarised perpendicular to the electron's orbit vanishes since $\hat{f}_\perp(\omega) = 0$. In the orbital plane, synchrotron radiation is thus completely linearly polarised in the orbital plane, or perpendicular to the guiding magnetic field. Then, with $\theta = 0$, (2.86) simplifies to

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} = \frac{3}{\pi}\frac{e^2\gamma^2}{c}\left(\frac{\omega}{\omega_c}\right)^2 K_{2/3}^2\left(\frac{\omega}{\omega_c}\right) , \quad \omega_c = 3\omega_L\gamma^3 \tag{2.87}$$

where we have introduced the cutoff frequency

$$\omega_c = 3\omega_L\gamma^3 = \frac{3\gamma^2 eB}{mc} . \tag{2.88}$$

The Bessel function $K_{2/3}(\xi)$ follows a falling power law for $\xi \ll 1$ and drops approximately exponentially for $\xi \gg 1$,

$$K_{2/3}(\xi) \approx \begin{cases} \frac{1}{2}\Gamma\left(\frac{2}{3}\right)\left(\frac{\xi}{2}\right)^{-2/3} & (\xi \ll 1) \\ \frac{\sqrt{\pi}}{\sqrt{2\xi}}\,\mathrm{e}^{-\xi} & (\xi \gg 1) \end{cases} . \tag{2.89}$$

For small ξ, the synchrotron spectrum is thus a power law in frequency,

$$\frac{\mathrm{d}^2E}{\mathrm{d}\Omega\mathrm{d}\omega} \approx \frac{3\cdot 2^{4/3}}{4\pi}\Gamma^2\left(\frac{2}{3}\right)\frac{e^2\gamma^2}{c}\left(\frac{\omega}{\omega_c}\right)^{2/3} \tag{2.90}$$

Since $\gamma \gg 1$, the frequency range covered by this power-law behaviour is very wide. Only far above the Larmor frequency, the spectrum is cut off exponentially near the cutoff frequency ω_c. This is a direct consequence of the narrow radiation cone: During each orbit of the electron, its radiation is only received by the observer in a very short time interval. The Fourier transform of this time interval, however, corresponds to a wide frequency range, similar to the uncertainty principle in quantum mechanics. For frequencies near or above the cutoff frequency, the spectrum is approximated by

$$\frac{\mathrm{d}^2 E}{\mathrm{d}\Omega \mathrm{d}\omega} \approx \frac{3}{2} \frac{e^2 \gamma^2}{c} \left(\frac{\omega}{\omega_\mathrm{c}}\right) \exp\left(-2\frac{\omega}{\omega_\mathrm{c}}\right) . \tag{2.91}$$

Problems

1. Due to the Lorentz force, a non-relativistic electron moving with a velocity v through the magnetic field $\vec{B}$ experiences the acceleration

$$\ddot{x} = -\frac{e}{mc}\left(\vec{v} \times \vec{B}\right) . \tag{2.92}$$

   (a) What is the average amount of energy per unit time and volume, $\mathrm{d}^2 E/(\mathrm{d}t\mathrm{d}V)$, radiated away by an isotropic electron distribution with number density n_e?

   (b) Assume now further that the electrons are in thermal equilibrium. In this case, the probability for an electron to have the velocity $v = |\vec{v}|$ is given by the Maxwell-Boltzmann distribution

$$p(v)\mathrm{d}v = \sqrt{\frac{2}{\pi}} \left(\frac{m_\mathrm{e}}{k_\mathrm{B}T}\right)^{3/2} v^2 \exp\left(-\frac{m_\mathrm{e} v^2}{2k_\mathrm{B}T}\right) , \tag{2.93}$$

   where T is the temperature of the electron gas, k_B is Boltzmann's constant and m_e the electron mass. Calculate $\mathrm{d}^2 E/(\mathrm{d}t\mathrm{d}V)$ as a function of the electron temperature T and the magnetic field $\vec{B}$. *Hint*: You can use that

$$\int_0^\infty \mathrm{d}x\, x^4 \mathrm{e}^{-ax^2} = \frac{3\sqrt{\pi}}{8} a^{-5/2} . \tag{2.94}$$

2. The synchrotron spectrum in the orbital plane of a single electron with Larmor frequency ω_L is

$$\frac{\mathrm{d}^2 E}{\mathrm{d}\omega \mathrm{d}\Omega} = \frac{3e^2\gamma^2}{\pi c} \left(\frac{\omega}{\omega_\mathrm{c}}\right)^2 K_{2/3}^2 \left(\frac{\omega}{\omega_\mathrm{c}}\right) , \tag{2.95}$$

   where $\omega_\mathrm{c} = 3\omega_\mathrm{L}\gamma^3$ and $K_{2/3}(x)$ is the modified Bessel function of order $2/3$ of the second kind.

   (a) In stochastic particle-acceleration processes, the accelerated electrons typically follow an energy distribution of the power-law form

$$\frac{\mathrm{d}N}{\mathrm{d}E}\mathrm{d}E = AE^{-\alpha}\mathrm{d}E , \tag{2.96}$$

   where A is a normalisation constant. Calculate the spectrum for such a population of electrons. *Hint*: Express the energy E by γ and and use

$$\int_0^\infty \mathrm{d}x\, x^a K_{2/3}^2(bx^2) = b^{-(a+1)/2} \frac{\sqrt{\pi}\Gamma\left(\frac{3a-5}{12}\right)\Gamma\left(\frac{3a+11}{12}\right)\Gamma\left(\frac{a+1}{4}\right)}{8\Gamma\left(\frac{a+3}{4}\right)} , \tag{2.97}$$

   valid for $a > 5/3$.

   (b) Draw the expected spectrum schematically in a double-logarithmic plot.

2.4 Bremsstrahlung

This section is concerned with a conceptually simple, but mathematically involved problem: Electrons scattering off ions follow hyperbolic orbits, are accelerated accordingly and emit free-free radiation or bremsstrahlung. A thermal ensemble of such electrons emits a spectrum characterised by an exponential cut-off, reflecting the Boltzmann factor of their energy distribution. The mathematical difficulty arises because, as we have seen in our general derivation of electromagnetic spectra, the hyperbolic electron orbits appears in the phase of a Fourier transform. This gives rise to Hankel functions of continuous order, which are difficult to handle. The main results of this section are the bremsstrahlung spectrum (2.118) of a single electron, the mean bremsstrahlung spectrum (2.122) after integrating over electron impact parameters, and the bremsstrahlung emissivity (2.131) obtained after integrating over a thermal electron population.

2.4.1 Orbit of an electron scattering off an ion

As we have seen before in (2.42), the spectrum of a non-relativistically moving charge is determined by the Fourier transform of its orbit $\vec{x}(t)$. Classically, an electron coming from infinity, scattering off an ion with charge Ze and leaving to infinity describes a hyperbolic orbit, much like a comet in the Solar System (Figure 2.7). We borrow the description of the orbit from the treatment of Kepler's problem in classical mechanics. By angular-momentum conservation, the orbit will be confined to a plane, in which we introduce plane polar coordinates (r, φ).

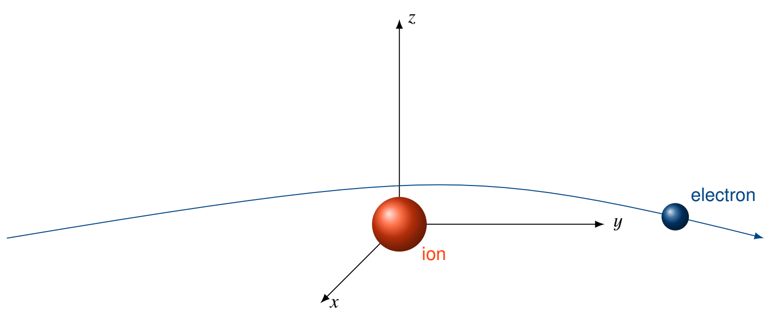

Figure 2.7 On the origin of bremsstrahlung: An electron is accelerated by the Coulomb force of an ion. It performs a hyperbolic orbit around the ion.

The (positive) energy of the electron is

$$E = \frac{m}{2}\dot{r}^2 + \frac{l^2}{2mr^2} - \frac{Ze^2}{r} , \tag{2.98}$$

where l is the conserved angular momentum. The solution of Kepler's problem tells us that this equation is solved by the conical sections (Figure 2.8), described in polar coordinates by

$$r(\varphi) = \frac{p}{1 + \varepsilon \cos \varphi} , \tag{2.99}$$

with the orbital parameter p and the numerical eccentricity ε expressing the angular momentum and the energy,

$$p = \frac{l^2}{Ze^2 m} \;, \quad \varepsilon^2 = 1 + \frac{2Ep}{Ze^2} \;. \tag{2.100}$$

Since e^2 must have the dimension erg cm in the Gaussian cgs system, it is quite easy to convince oneself that p is a length and ε is dimension-less. We further introduce the length scale a by

$$p = a(\varepsilon^2 - 1) \;. \tag{2.101}$$

For a bound elliptical orbit, a is the semi-major axis. Combining (2.101) with the second equation (2.100), we can express the energy by the orbital parameter a as

$$E = \frac{Ze^2}{2a} \;. \tag{2.102}$$

?

If needed, recapitulate the derivation of equation (2.99) for Kepler orbits, and the conditions for it to be valid.

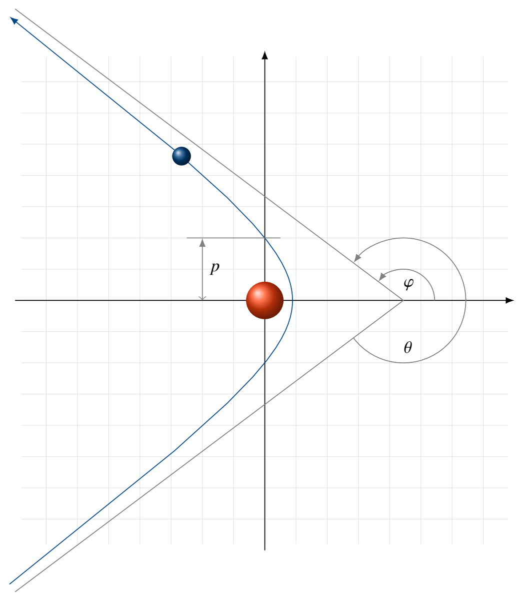

Figure 2.8 Hyperbolic orbit of an unbound particle in an attractive field of force.

We now replace the polar angle φ by the so-called *eccentric anomaly* ψ. For an unbound orbit, ψ is implicitly defined by

$$r(\psi) = a(\varepsilon \cosh\psi - 1) \;, \tag{2.103}$$

which, together with (2.99) and (2.101) implies

$$\cos\varphi = \frac{\varepsilon - \cosh\psi}{\varepsilon\cosh\psi - 1} \;. \tag{2.104}$$

Now, we eliminate the squared angular momentum l^2 between (2.100) and (2.98), insert the expression (2.102) for the energy into the resulting equation and solve it for $\dot{r}^2$,

$$\dot{r}^2 = \frac{2}{m}\left[\frac{Ze^2}{2a} + \frac{Ze^2}{r} - \frac{Ze^2 a(\varepsilon^2 - 1)}{2r^2}\right] = \frac{2Ze^2}{mr^2}\left[\frac{r^2}{2a} + r - \frac{a(\varepsilon^2 - 1)}{2}\right] \;. \tag{2.105}$$

Next, we use (2.103) to introduce the eccentric anomaly into the following terms,

$$\dot{r} = a\varepsilon \sinh\psi\dot{\psi} \; , \quad \frac{r^2}{2a} + r - \frac{a(\varepsilon^2 - 1)}{2} = \frac{a\varepsilon^2}{2}\sinh^2\psi \; . \tag{2.106}$$

These finally allow us to bring (2.105) into the form

$$\dot{\psi}^2 = \frac{Ze^2}{ma^3(\varepsilon\cosh\psi - 1)^2} \; . \tag{2.107}$$

Separating the variables ψ and t, we can express the time needed by the particle to get from $\psi = 0$ to ψ as

$$t = \int_0^t \mathrm{d}t' = \sqrt{\frac{ma^3}{Ze^2}} \int_0^\psi \mathrm{d}\psi'(\varepsilon\cosh\psi' - 1) = \tau(\varepsilon\sinh\psi - \psi) \; , \tag{2.108}$$

where the time scale τ was introduced. Equation (2.108) is *Kepler's equation* for a hyperbolic orbit.

Caution Note that Kepler's equation is transcendental and can thus only be solved numerically. ◀

Since the energy E must be the kinetic energy of the electron at infinite distance from the ion, we can eliminate a from (2.102),

$$\frac{m}{2}v_\infty^2 = E = \frac{Ze^2}{2a} \quad \Rightarrow \quad a = \frac{Ze^2}{mv_\infty^2} \; . \tag{2.109}$$

In terms of v_∞, the time scale τ is thus given by

$$\tau = \sqrt{\frac{ma^3}{Ze^2}} = \frac{Ze^2}{mv_\infty^3} = \frac{a}{v_\infty} \; . \tag{2.110}$$

2.4.2 Fourier transform of the orbit

By means of Kepler's equation (2.108), we can now substitute the time t by the eccentric anomaly ψ in the Fourier transform of the electron's orbit. First, we combine (2.103) and (2.104) to write the Cartesian coordinates

$$\begin{aligned} x(\psi) &= r\cos\phi = a(\varepsilon - \cosh\psi) \; , \\ y(\psi) &= \sqrt{r^2 - x^2} = a\sqrt{\varepsilon^2 - 1}\sinh\psi \; . \end{aligned} \tag{2.111}$$

Moreover, we have from (2.108)

$$\mathrm{d}t = \tau(\varepsilon\cosh\psi - 1)\mathrm{d}\psi \; , \quad \mathrm{e}^{\mathrm{i}\omega t} = \mathrm{e}^{\mathrm{i}\omega\tau(\varepsilon\sinh\psi - \psi)} \; . \tag{2.112}$$

It is now convenient to compute the Fourier transform of the velocity, $\hat{\vec{v}} = -\mathrm{i}\omega\hat{\vec{x}}$, instead of the Fourier transform of the orbit, $\hat{\vec{x}}$. We thus write

$$\begin{aligned} \hat{x}(\omega) &= -\frac{\hat{\dot{x}}}{\mathrm{i}\omega} \int_{-\infty}^{\infty} \mathrm{d}t\,\dot{x}\,\mathrm{e}^{-\mathrm{i}\omega t} = \frac{\mathrm{i}}{\omega}\int_{-\infty}^{\infty}\mathrm{d}t\,\frac{\mathrm{d}x}{\mathrm{d}\psi}\,\dot{\psi}\,\mathrm{e}^{-\mathrm{i}\omega t} \\ &= \frac{\mathrm{i}}{\omega}\int_{-\infty}^{\infty}\mathrm{d}\psi\,\frac{\mathrm{d}x}{\mathrm{d}\psi}\,\mathrm{e}^{-\mathrm{i}\omega t(\psi)} \; , \end{aligned} \tag{2.113}$$

and likewise for $\hat{y}(\omega)$. With (2.111), this gives

$$\hat{x}(\omega) = -\frac{\mathrm{i}a}{\omega}\int_{-\infty}^{\infty}\mathrm{d}\psi\,\sinh\psi\,\mathrm{e}^{-\mathrm{i}\omega\tau(\varepsilon\sinh\psi-\psi)}\,,$$
$$\hat{y}(\omega) = \frac{\mathrm{i}a\sqrt{\varepsilon^2-1}}{\omega}\int_{-\infty}^{\infty}\mathrm{d}\psi\,\cosh\psi\,\mathrm{e}^{-\mathrm{i}\omega\tau(\varepsilon\sinh\psi-\psi)}\,. \tag{2.114}$$

These integrals can be expressed by the Hankel function of the first kind of order ν, $H_\nu^{(1)}(x)$, and its derivative, $H_\nu^{(1)\prime}(x)$. In terms of these, we have

$$\hat{x}(\omega) = \frac{\pi a}{\omega}H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu\varepsilon)\,,\quad \hat{y}(\omega) = -\frac{\pi a\sqrt{\varepsilon^2-1}}{\omega\varepsilon}H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu\varepsilon)\,, \tag{2.115}$$

where the order $\nu = \omega\tau$. With (2.42), we thus find the bremsstrahlung spectrum

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{2\pi^2a^2e^2\omega^2}{3c^3}\left\{\left[H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu\varepsilon)\right]^2-\left(1-\frac{1}{\varepsilon^2}\right)\left[H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu\varepsilon)\right]^2\right\} \tag{2.116}$$

for a single electron moving on a hyperbolic orbit with eccentricity ε. The sign in front of the second term in brackets is negative because $H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu\varepsilon)$ is purely imaginary, while its derivative $H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu\varepsilon)$ is real. Before we can continue, we need to integrate (2.116) over a realistic distribution of the eccentricity ε.

Caution The Hankel function of the first kind is the complex linear combination

$$H_\nu^{(1)}(x) = J_\nu(x) + \mathrm{i}Y_\nu(x)$$

of the Bessel functions J_ν and Y_ν of the first and second kinds. Both solve Bessel's differential equation

$$x^2\frac{\mathrm{d}^2f}{\mathrm{d}x^2} + x\frac{\mathrm{d}f}{\mathrm{d}x} + \left(x^2-\nu^2\right)f = 0\,.$$

◀

The following relation between Bessel functions and their derivatives comes to help, which also applies to the Hankel functions,

$$z\left[Z_p^{\prime 2}(z) - \left(1-\frac{p^2}{z^2}\right)Z_p^2(z)\right] = \frac{\mathrm{d}}{\mathrm{d}z}\left(zZ_p(z)Z_p'(z)\right)\,. \tag{2.117}$$

Setting $z = \mathrm{i}\nu\varepsilon$ and $p = \mathrm{i}\nu$, this allows us to write (2.116) as

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = -\mathrm{i}\,\frac{2\pi^2a^2e^2\omega}{3\tau\varepsilon c^3}\,\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\varepsilon H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu\varepsilon)H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu\varepsilon)\right]\,, \tag{2.118}$$

where we have used that the order $\nu = \omega\tau$. The prefactor $-\mathrm{i}$ is necessary because the Hankel function $H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu\varepsilon)$ is imaginary.

2.4.3 Integration over impact parameters

The numerical eccentricity ε of an particle's orbit is determined by its angular momentum l, which is in turn controlled by the orbit's impact parameter b. This is defined as the closest distance of the scattering centre from the straight line which would be the electron's unperturbed trajectory. Combining the two equations (2.100) with $E = mv_\infty^2/2$, we find

$$\varepsilon^2 = 1 + \frac{v_\infty^2l^2}{Z^2e^4} \tag{2.119}$$

for the squared numerical eccentricity. Then, using the expression $l = bmv_\infty$ for the angular momentum and replacing the constants occurring by means of (2.109), the simple result is

$$\varepsilon^2 = 1 + \frac{b^2}{a^2}\,. \tag{2.120}$$

We now take the spectrum (2.118) produced by a single electron and multiply it with the number of scattering events between electrons and ions per unit time and unit volume. Let n_i and n_e be the number densities of ions and electrons, respectively, and v_∞ the velocity of the electrons relative to the ions. Consider a single ion and surround it by a cylindrical shell of radius b, width $\mathrm{d}b$, and height $v_\infty \mathrm{d}t$. Then, all

$$n_\mathrm{e} \cdot (2\pi b \mathrm{d}b) \cdot (v_\infty \mathrm{d}t) = 2\pi n_\mathrm{e} v_\infty a^2 \, \varepsilon \mathrm{d}\varepsilon \, \mathrm{d}t \tag{2.121}$$

electrons contained in this shell will scatter off the ion within the time interval $\mathrm{d}t$. Multiplying this number with n_i, we find the total number of scatterings between ions and electrons with relative velocity v_∞ and impact parameter within $[b, b + \mathrm{d}b]$ per unit time and unit volume. Further multiplying this number with the spectrum (2.118), and integrating over all impact parameters b or eccentricities ε, then gives the spectrum emitted by such electrons per unit time and volume,

$$\frac{\mathrm{d}^3 E}{\mathrm{d}\omega \mathrm{d}t \mathrm{d}V} = \mathrm{i}\, \frac{4\pi^3 Z^2 e^6 n_\mathrm{i} n_\mathrm{e}}{3m^2 c^3 v_\infty} \left(\frac{Ze^2\omega}{mv_\infty^3}\right) H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu) H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu) \ . \tag{2.122}$$

For arriving at this expression, we have used (2.109) and (2.110) to substitute a and τ and regrouped terms for later convenience.

The Hankel functions and their derivatives need to be numerically evaluated, but we can insert their asymptotic forms for small and large arguments. These are

$$\begin{aligned} \nu \ll 1 : H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu) &\approx \frac{2}{\mathrm{i}\pi} \ln\left(\frac{2}{\gamma\nu}\right) , \quad H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu) \approx \frac{2}{\pi\nu} \\ \nu \gg 1 : H_{\mathrm{i}\nu}^{(1)}(\mathrm{i}\nu) &\approx -\frac{\mathrm{i}}{\pi\sqrt{3}} \left(\frac{6}{\nu}\right)^{1/3} \Gamma(1/3) , \quad H_{\mathrm{i}\nu}^{(1)\prime}(\mathrm{i}\nu) \approx \frac{1}{\pi\sqrt{3}} \left(\frac{6}{\nu}\right)^{2/3} \Gamma(2/3) . \end{aligned} \tag{2.123}$$

Now, with the further help of

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x} , \quad \Gamma(1/3)\Gamma(2/3) = \frac{\pi}{\sin(\pi/3)} = \frac{2\pi}{\sqrt{3}} , \tag{2.124}$$

we find the low- and high-frequency approximations

$$\frac{\mathrm{d}^3 E}{\mathrm{d}\omega \mathrm{d}t \mathrm{d}V} = \frac{16\pi Z^2 e^6 n_\mathrm{i} n_\mathrm{e}}{3m^2 c^3 v_\infty} \begin{cases} \ln\left(\dfrac{2}{\gamma} \dfrac{mv_\infty^3}{Ze^2\omega}\right) & \omega \ll \tau^{-1} \\ \dfrac{\pi}{\sqrt{3}} & \omega \gg \tau^{-1} \end{cases} . \tag{2.125}$$

Recall from (2.110) that $\tau = a v_\infty^{-1}$

2.4.4 Average over electron velocities, thermal bremsstrahlung

The dependence of the spectrum on ω is mild for low ω, and absent for high ω, which is a very interesting result: The energy emitted per unit frequency is (almost) independent of the frequency. These asymptotic results motivate writing the complete spectrum of non-relativistic bremsstrahlung in the form

$$\frac{\mathrm{d}^3E}{\mathrm{d}\omega\mathrm{d}t\mathrm{d}V} = j(\omega) = \frac{16\pi^2Z^2e^6n_\mathrm{i}n_\mathrm{e}}{3\sqrt{3}m^2c^3}\frac{g_\mathrm{ff}(v_\infty,\omega)}{v_\infty}\,, \tag{2.126}$$

introducing the so-called Gaunt factor $g_\mathrm{ff}(v_\infty,\omega)$. In the high-frequency limit, g_ff tends to unity, as (2.125) shows, and depends generally only weakly on v_∞ and ω. It is thus reasonable to introduce a velocity-averaged Gaunt factor by

$$\left\langle\frac{g_\mathrm{ff}(v_\infty,\omega)}{v_\infty}\right\rangle = \bar{g}_\mathrm{ff}(\omega)\left\langle\frac{1}{v_\infty}\right\rangle \tag{2.127}$$

and average the reciprocal velocity over some velocity distribution.

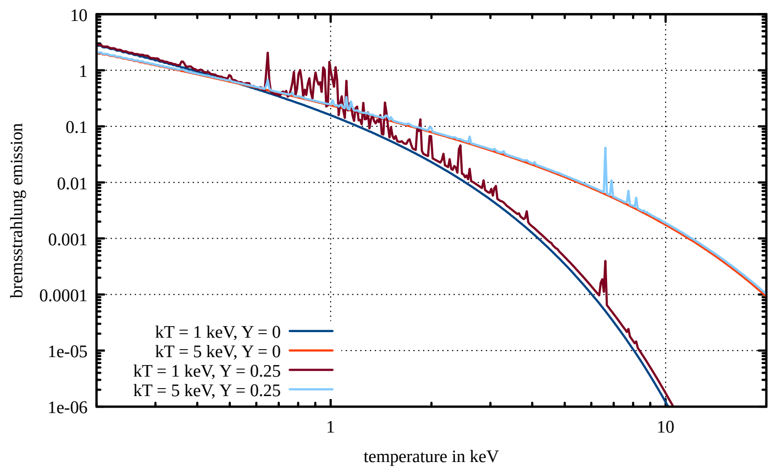

Figure 2.9 Thermal bremsstrahlung without and with line emission, for plasma temperatures of 1 keV and 5 keV. The spectra were produced with the `xspec` software package using a Raymond-Smith plasma model.

If the electrons scattering off the ions form a thermal population, their velocity distribution is Maxwellian,

$$p(v_\infty)\mathrm{d}v_\infty = 4\pi\left(\frac{m}{2\pi k_\mathrm{B}T}\right)^{3/2} v_\infty^2 \exp\left(-\frac{mv_\infty^2}{2k_\mathrm{B}T}\right)\mathrm{d}v_\infty\,. \tag{2.128}$$

For emitting at least a single photon of frequency ω or energy $\hbar\omega$, an electron has to satisfy

$$\frac{mv_\infty^2}{2} \ge \hbar\omega \quad\Rightarrow\quad v_\infty \ge v_\mathrm{min} = \sqrt{\frac{2\hbar\omega}{m}}\,. \tag{2.129}$$

The average of v_∞^{-1} then turns out to be

$$\left\langle\frac{1}{v_\infty}\right\rangle = 4\pi\left(\frac{m}{2\pi k_\mathrm{B}T}\right)^{3/2}\int_{v_\mathrm{min}}^\infty v_\infty\mathrm{d}v_\infty \exp\left(-\frac{mv_\infty^2}{2k_\mathrm{B}T}\right) = \sqrt{\frac{2m}{\pi k_\mathrm{B}T}}\exp\left(-\frac{\hbar\omega}{k_\mathrm{B}T}\right)\,. \tag{2.130}$$

Combined with (2.126), this finally gives the emissivity of non-relativistic, thermal bremsstrahlung (Figure 2.9)

$$j(\omega) = \frac{16\pi^2}{3\sqrt{3}}\frac{Z^2 e^6 n_\mathrm{i} n_\mathrm{e}}{m^2 c^3}\bar{g}_\mathrm{ff}(\omega)\sqrt{\frac{2m}{\pi k_\mathrm{B} T}}\exp\left(-\frac{\hbar\omega}{k_\mathrm{B} T}\right) . \tag{2.131}$$

The Gaunt factor is typically tabulated, but for many astrophysical applications, $\bar{g}_\mathrm{ff}(\omega) \approx 1$ is a sufficient approximation.

Problems

1. A simplified derivation of the bremsstrahlung emissivity begins with Born's approximation, asserting that the electron's acceleration can be evaluated along a straight, undeflected orbit.

   (a) Evaluate the electron's acceleration by an ion along a straight line.

   (b) Fourier transform the acceleration and calculate the approximate bremsstrahlung spectrum.

   (c) Carry out the integration over impact parameters. Which problem occurs?

2.5 Radiation damping

Remarkably, electrodynamics is incomplete in the following sense: Consider an electron moving in a homogeneous magnetic field in the absence of electric fields. The Lorentz force then causes the electron to move on a spiral orbit without changing the electron's energy. At that level, the prediction of electrodynamics would be that the electron keeps moving in this way forever. However, the motion along the spiral is an accelerated motion, which implies that the electron loses energy by radiation. This loss of energy is not contained in the equation of motion for the electron. The back-reaction of the radiation emitted by an accelerated charge on the motion of that same charge has to be described separately. This is a fundamental limit of electrodynamics: As a linear theory, it cannot encompass this kind of back-reaction. In this section, the backreaction of the radiation on the radiating charge itself is derived. The loss of energy by the charge due to the radiation can be described by an effective force, the radiation-damping force, the expression (2.139) for which will be the first main result. As an important application, the energy transfer from a charge moving through a sea of radiation to that radiation field itself is developed next, which leads to the very intuitive result (2.163) for the transferred power.

2.5.1 Damping force

The loss of energy by radiation can be described as the action of an effective damping force $\vec{F}_{\rm rad}$ acting on the electron. The energy radiated away within a certain time interval $-\tau/2 \le t \le \tau/2$,

$$E = \int_{-\tau/2}^{\tau/2} \mathrm{d}t\, P(t) \;, \tag{2.132}$$

P being the radiative power, must then equal the work exerted by this radiation-damping force on the electron during the same time,

$$\int_{-\tau/2}^{\tau/2} \mathrm{d}t\, P(t) = -\int \vec{F}_{\rm rad} \cdot \mathrm{d}\vec{s} \;. \tag{2.133}$$

The solid-angle integrated Larmor formula (1.141) shows that the power radiated by an accelerated electron is homogeneous of degree $k = 2$ in the acceleration $\dot{\beta}$, that is, if the acceleration is scaled by a dimension-less factor a, the power changes by a factor a^2. Generally, a function $f(x)$ is called homogeneous of degree k if $f(ax) = a^k f(x)$ for $a \in \mathbb{R}$. The Larmor power thus satisfies Euler's theorem for homogeneous functions: If $f(\vec{x})$ is a homogeneous function of degree k in $\vec{x}$, then its derivative satisfies

$$\vec{x} \cdot \frac{\mathrm{d}f(\vec{x})}{\mathrm{d}\vec{x}} = k f(\vec{x}) \;. \tag{2.134}$$

?

Can you prove Euler's theorem (2.134) for homogeneous functions? Otherwise, look it up.

When applied to the radiation power, Euler's theorem thus says

$$\dot{\vec{\beta}} \cdot \frac{\partial}{\partial \dot{\vec{\beta}}} P\left(\dot{\vec{\beta}}\right) = 2P\left(\dot{\vec{\beta}}\right) \;. \tag{2.135}$$

We use this statement to express the power in (2.133) by its derivative and obtain

$$-\int \vec{F}_{\rm rad} \cdot \mathrm{d}\vec{s} = \frac{1}{2} \int \mathrm{d}t\, \dot{\vec{\beta}} \cdot \frac{\partial}{\partial \dot{\vec{\beta}}} P\left(\dot{\vec{\beta}}\right) = -\frac{1}{2} \int \mathrm{d}t\, \vec{\beta} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{\vec{\beta}}} P\left(\dot{\vec{\beta}}\right) \tag{2.136}$$

by partial integration, omitting the boundary terms. This is generally no substantial restriction because we can typically choose the integration boundaries wide enough for the radiation power to vanish at both of them. Now, since $\vec{\beta}\mathrm{d}t = \mathrm{d}\vec{s}/c$, we can identify the expression

$$\vec{F}_{\rm rad} = \frac{1}{2c} \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{\vec{\beta}}} P\left(\dot{\vec{\beta}}\right) \tag{2.137}$$

with the radiation-damping force. In the non-relativistic limit (1.142),

$$P = \frac{2e^2}{3c} \dot{\vec{\beta}}^2 \;, \quad \frac{\partial}{\partial \dot{\vec{\beta}}} P\left(\dot{\vec{\beta}}\right) = \frac{4e^2}{3c} \dot{\vec{\beta}} \;, \tag{2.138}$$

whence the radiation-damping force turns out to be

$$\vec{F}_{\rm rad} = \frac{2e^2}{3c^2} \ddot{\vec{\beta}} \tag{2.139}$$

Example: Scattering off bound electrons

We will now directly apply this result to an electron on a bound harmonic orbit with an angular frequency ω_0. Let the electron be externally driven by an incoming electromagnetic wave with frequency ω. This wave exerts the electric Lorentz force

$$\vec{F}_\mathrm{L} = -\frac{e}{c}\,\vec{E}_0\,\mathrm{e}^{\mathrm{i}\omega t} \tag{2.140}$$

on the electron. We assume that the electron moves non-relativistically such that we can ignore the magnetic part of the Lorentz force. Including radiation damping with a damping constant γ to be determined shortly, the equation of motion

$$\ddot{\vec{x}} + \gamma\dot{\vec{x}} + \omega_0^2\vec{x} = -\frac{e}{m}\,\vec{E}_0\,\mathrm{e}^{\mathrm{i}\omega t} \tag{2.141}$$

describes a harmonically driven and damped harmonic oscillator. Its particular solution is immediately found to read

$$\vec{x} = -\frac{e}{m}\,\frac{\vec{E}_0\,\mathrm{e}^{\mathrm{i}\omega t}}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma} \tag{2.142}$$

after an initial settling phase during which a possible oscillation with the eigenfrequency ω_0 of the bound orbit decays exponentially. We thus have

$$\ddot{\vec{\beta}} = -\omega^2\vec{\beta}\,, \tag{2.143}$$

allowing us to write the radiation-damping force as

$$\vec{F}_\mathrm{rad} = -\frac{2e^2\omega^2}{3c^2}\,\vec{\beta} \tag{2.144}$$

and to identify the damping constant

$$\gamma = \gamma_0\omega^2 \quad \text{with} \quad \gamma_0 = \frac{2}{3}\frac{e^2}{mc^3} = \frac{2}{3}\frac{r_\mathrm{e}}{c}\,, \tag{2.145}$$

where r_e is the classical electron radius introduced in (2.12). According to (2.142), the electron's acceleration is

$$\dot{\vec{\beta}} = -\omega^2\,\frac{\vec{x}}{c} = \frac{e}{mc}\,\frac{\vec{E}_0\,\mathrm{e}^{\mathrm{i}\omega t}\,\omega^2}{\omega_0^2 - \omega^2 - \mathrm{i}\gamma_0\omega^3}\,, \tag{2.146}$$

which we can now insert into the non-relativistic, integrated Larmor equation (1.142) to find

$$P = \frac{2e^2}{3c}\left|\dot{\vec{\beta}}\right|^2 = \frac{2e^4}{3m^2c^3}\,\vec{E}_0^2\,\frac{\omega^4}{\left(\omega^2 - \omega_0^2\right)^2 + \gamma_0^2\omega^6}\,. \tag{2.147}$$

The incoming energy current density is given by the amplitude of the Poynting vector $|\vec{S}| = c\vec{E}_0^2/(4\pi)$, and thus the cross section for scattering off a harmonically bound charge becomes

$$\sigma = \frac{P}{\left|\vec{S}\right|} = \sigma_\mathrm{T}\,\frac{\omega^4}{\left(\omega^2 - \omega_0^2\right)^2 + \gamma_0^2\omega^6} \tag{2.148}$$

with the typical resonance behaviour near $\omega = \omega_0$ (Figure 2.10). ◂

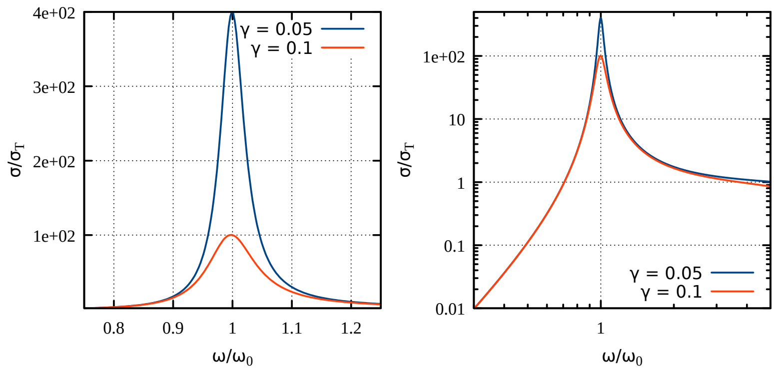

Figure 2.10 Illustration of the cross section for scattering of electromagnetic radiation off a harmonically bound electron. *Left*: The cross section (2.148) is shown (in units of the Thomson cross section σ_T) for two values of the damping constant. *Right*: The same curves, now plotted double-logarithmically, reveal the ω^4 scaling for low frequencies, i.e. the regime of Rayleigh scattering.

in this limit. Since $\vec\beta = \dot{\vec x}/c$, this involves a third time derivative of the electron's orbit. This is one of the rare cases of a third-order time derivative in physics.

Some limiting cases of the general cross section (2.148) for scattering off bound electrons are of particular interest. First, in the high-frequency limit $\omega \gg \omega_0$ and $\omega \gg \gamma_0^{-1}$, the driving force oscillates so fast that radiation damping is strong. The cross section (2.148) then falls off like ω^{-2},

$$\sigma \approx \frac{\sigma_\mathrm{T}}{\gamma_0^2\omega^2}\;. \tag{2.149}$$

?

Verify the combined results (2.152) and (2.153) and confirm that the Lorentz profile is normalised to unity.

Notice, however, that the electron will be unbound if the incoming radiation has too high frequency, and then its cross section will turn into the Thomson cross section, $\sigma \approx \sigma_\mathrm{T}$.

In the opposite limit, when $\omega \ll \omega_0$ and $\omega \ll \gamma_0^{-1}$, we find the limit of Rayleigh scattering,

$$\sigma \approx \sigma_\mathrm{T}\left(\frac{\omega}{\omega_0}\right)^4\;, \tag{2.150}$$

with the scattering cross section depending on the fourth power of the frequency. For $\omega \approx \omega_0$ and weak damping, $\omega_0 \ll \gamma_0^{-1}$, we approximate

$$\omega^2 - \omega_0^2 = (\omega - \omega_0)(\omega + \omega_0) \approx 2\omega_0(\omega - \omega_0) \tag{2.151}$$

in (2.148) and find

$$\sigma \approx \frac{\pi}{2}\frac{\sigma_\mathrm{T}}{\gamma_0}\,\phi_\Gamma\,(\omega - \omega_0)\;,\quad \Gamma := \gamma_0\omega_0^2\;, \tag{2.152}$$

where the function $\phi_\Gamma(\omega - \omega_0)$ is the so-called *Lorentz profile*,

$$\phi_\Gamma(\omega - \omega_0) = \frac{1}{\pi}\frac{\Gamma/2}{(\omega - \omega_0)^2 + (\Gamma/2)^2} \tag{2.153}$$

shown in Fig. 2.11. The Lorentz profile will recur several times in later Sections. It is normalised to unity.

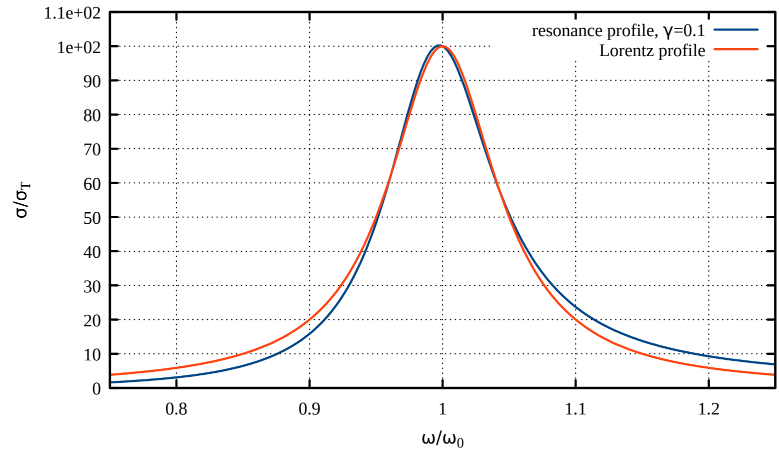

Figure 2.11 Near the resonance, the scattering cross section is reasonably approximated by the Lorentz profile.

2.5.2 Transfer of energy from a moving charge to a radiation field

Consider now an electron moving with possibly relativistic speed $\vec{\beta}$ through an isotropic radiation field, whose electric and magnetic field components satisfy

$$\left\langle \vec{E} \right\rangle = 0 = \left\langle \vec{B} \right\rangle \;, \tag{2.154}$$

where the average is taken over time intervals long compared to typical oscillation period $2\pi\omega^{-1}$ of the radiation field. Now we transform to the rest frame of the electron. The electron experiences the field components $\vec{E}'$, $\vec{B}'$ given by the Lorentz transform (1.87). They accelerate the electron through the electric Lorentz force

$$\ddot{\vec{x}}' = \frac{1}{m}\vec{F}'_{\rm L} = \frac{e}{m}\vec{E}' \tag{2.155}$$

since the magnetic part of the Lorentz force vanishes in the electron's rest frame, where $\vec{v}' = 0$. We can now calculate the power radiated by the accelerated electron with the non-relativistic Larmor formula, for which we need to evaluate

$$\left\langle \left|\ddot{\vec{x}}'\right|^2 \right\rangle = \frac{e^2}{m^2}\left\langle \left|\vec{E}'\right|^2 \right\rangle \tag{2.156}$$

in the electron's rest frame. Here, we can directly insert the Lorentz transform of the fields from (1.87) and carry out the average. Doing so, we have to take into account that the electromagnetic field in its rest frame is randomly oriented and has an energy density U. This allows us to use

$$\left\langle E_i^2 \right\rangle = \frac{\left\langle \vec{E}^2 \right\rangle}{3} = \frac{4\pi}{3}U = \left\langle B_j^2 \right\rangle \tag{2.157}$$

for the squares of the electric and magnetic field components and

$$\left\langle E_i B_j \right\rangle = 0 \tag{2.158}$$

for any combination of i and j. These relations enable us to write

$$\left\langle \left|\ddot{\vec{x}}'\right|^2 \right\rangle = 4\pi\gamma^2 U\frac{e^2}{m^2}\left(\frac{2}{3} + \frac{1}{3\gamma^2} + \frac{2}{3}\beta^2\right) = 4\pi\gamma^2 U\frac{e^2}{m^2}\left(1 + \frac{\beta^2}{3}\right) \tag{2.159}$$

and thus

$$P_{\rm em} = \gamma^2 U \frac{8\pi e^4}{3m^2c^3}\left(1+\frac{\beta^2}{3}\right) = cU\sigma_{\rm T}\gamma^2\left(1+\frac{\beta^2}{3}\right) \tag{2.160}$$

for the power radiated by the electron in its rest frame. However, since the power is

$$P = \frac{\mathrm{d}E}{\mathrm{d}t} \tag{2.161}$$

and both the energy E and the time t transform like the zero components of four-vectors, the power is invariant under Lorentz transforms. Therefore, the result (2.160) also holds in the rest frame of the radiation field. On the other hand, the power absorbed by the electron is given by the Poynting vector times the cross section,

$$P_{\rm abs} = \left|\vec{S}\right|\sigma_{\rm T} = \frac{c}{4\pi}\vec{E}^2\sigma_{\rm T} = cU\sigma_{\rm T}\ . \tag{2.162}$$

The net power transferred by the electron to the radiation field is thus

$$P = P_{\rm em} - P_{\rm abs} = cU\sigma_{\rm T}\left[\gamma^2\left(1+\frac{\beta^2}{3}\right)-1\right] = \frac{4}{3}\beta^2\gamma^2 cU\sigma_{\rm T}\ . \tag{2.163}$$

?

Carrying out the description following (2.156), verify the expressions (2.159) and (2.160) by your own calculation.

We can now proceed to calculate the back-reaction on the electron by its transfer of energy to the radiation field. Clearly, the loss of kinetic energy of the electron must equal the negative radiation power (2.163),

$$\frac{\mathrm{d}E}{\mathrm{d}t} = mc^2\frac{\mathrm{d}\gamma}{\mathrm{d}t} = -\frac{4}{3}\beta^2\gamma^2 cU\sigma_{\rm T} = -\frac{4}{3}\left(\gamma^2-1\right)cU\sigma_{\rm T}\ . \tag{2.164}$$

Separating the variables γ and t and integrating over time gives

$$\int_\gamma^0 \frac{\mathrm{d}x}{x^2-1} = -\frac{t}{\tau} \quad \text{with} \quad \tau := \frac{3mc}{4U\sigma_{\rm T}}\ . \tag{2.165}$$

?

Explain the integral boundaries on the left-hand side of (2.165).

Noticing that

$$\frac{1}{x^2-1} = \frac{1}{2}\left(\frac{1}{x-1}-\frac{1}{x+1}\right)\ , \tag{2.166}$$

we can readily carry this integral out, finding

$$\frac{1}{2}\ln\frac{\gamma-1}{\gamma+1} = -\frac{t}{\tau}\ . \tag{2.167}$$

This equation can now easily be solved for γ or β, giving the essentially exponential decrease

$$\beta(t) = \frac{2\exp(-t/\tau)}{1+\exp(-2t/\tau)} \tag{2.168}$$

of the electron's velocity with time. This result shows that relativistic electrons, or charges in general, lose energy on a characteristic time scale

$$\tau = \frac{3mc}{4U\sigma_{\rm T}} \tag{2.169}$$

when interacting with a radiation field. As the expression shows, the time scale is given by the rest-mass energy of the electron, divided by the energy of the radiation field flowing per unit time through the Thomson cross section. Similarly, the characteristic path length for a relativistic electron to lose its energy in a radiation field is

$$\lambda = c\tau = \frac{3mc^2}{4U\sigma_{\rm T}}\ . \tag{2.170}$$

Problems

1. Derive the solution (2.145) of the equation of motion (2.143).

2. Calculate the time scale (2.169) for an electron travelling through the Cosmic Microwave Background.

3. The velocity of an electron in a homogeneous magnetic field changes due to the Lorentz force according to

$$\frac{\mathrm{d}\left(\gamma\vec{v}\right)}{\mathrm{d}t} = \frac{e}{m_{\mathrm{e}}c}\left(\vec{v}\times\vec{B}\right) \tag{2.171}$$

   (a) Set up the equations of motion for the individual components of $\vec{x}$ in the field $\vec{B} = B\hat{e}_z$.

   (b) How does the equation of motion change if the radiation damping force

$$\vec{F}_{\mathrm{rad}} = \frac{2e^2}{3c^3}\dddot{\vec{x}} \tag{2.172}$$

   is also taken into account? Assume that the energy loss per orbit is small, i.e. the damping force can be evaluated using the undamped solution from (a). Under which circumstances is the former assumption valid?

   (c) Solve the differential equations for the components x_i with the boundary conditions $\vec{x}(t = 0) = (x_0, 0, 0)^\top$ and $\vec{v}(t = 0) = (0, v_0, 0)^\top$. Draw the solution schematically.

2.6 Compton scattering

This section introduces the photon picture for electromagnetic radiation. So far, incoming electromagnetic waves could only accelerate charges perpendicular to their direction of motion, which implies that they could not transfer momentum to the charges. With the discussion of radiation damping in the preceding section, we have seen how charges experience an effective force against their direction of motion due to the radiation they emit. In the discussion of Compton scattering, the incoming radiation is described as a stream of photons transfering both energy and momentum to the charges they scatter off from. The main result derived here is the mean energy loss per photon per collision (2.183). We then proceed to calculating the energy gained by a moving charge from a sea of radiation by Compton scattering and combine it with the loss due to radiation damping to find the total rate (2.193) of energy transfer between the charge and the photons. Compton scattering is then combined with the Fokker-Planck approach to work out photon diffusion in phase space due to scattering with electrons. The main result there is the approximation (2.220) to the so-called Kompaneets equation which neglects effects from quantum statistics, but is nonetheless appropriate for many astrophysical circumstances.

2.6.1 Energy change in the scattering process

So far, we have studied how charges radiate when they are accelerated under several kinds of circumstances. We have seen in the last section how a charge can transfer energy to a radiation field by radiation damping.

Recall the physical situation we had in mind: A charge, say an electron, moving through an isotropic sea of radiation keeps being accelerated by the randomly oriented electromagnetic fields of the radiation sea. Due to this acceleration, the charge radiates away part of its kinetic energy and thus transfers energy to the radiation field.

Let us now consider the reverse question: Suppose we have an electron at rest and a radiation field streaming past it. Does the radiation field transfer any energy to the charge? In the classical picture of radiation being composed of electromagnetic waves, the charge is accelerated by the Lorentz force of the randomly superposed electromagnetic waves constituting the radiation field. The magnetic part of the Lorentz force can never change the charge's energy since it acts perpendicular to the charge's velocity.

Since electromagnetic waves in vacuum are transversal, the electric part of the Lorentz force cannot act in the streaming direction of the radiation in the charge's rest frame. Driven by the electric Lorentz force of the radiation, the charge will thus oscillate perpendicular to the streaming direction. If the radiation is unpolarised, the electric field experienced by the charge will be randomly superposed of waves with arbitrary orientations and random phases. Does this mean that there is no net energy transfer from the radiation field to the charge?

At this point, it is necessary to change to the photon picture and describe radiation as a stream of particles, each carrying a four-momentum

$$k^\mu = \frac{\omega}{c}\begin{pmatrix} 1 \\ \hat{e} \end{pmatrix} , \tag{2.173}$$

where $\hat{e}$ is the direction of motion. The total energy-momentum four-vector of the electron, p^μ, and the photon $\hbar k^\mu$ is conserved, and thus (Figure 2.12)

$$p^\mu + \hbar k^\mu = p'^\mu + \hbar k'^\mu , \tag{2.174}$$

where primes denote quantities after scattering. Recall the result (1.63) from relativistic dynamics, showing that the four-momentum of the electron has the components

$$p^\mu = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix} = \gamma m \begin{pmatrix} c \\ \vec{v} \end{pmatrix} \tag{2.175}$$

and the Minkowski square given by (1.65), $\langle p, p\rangle = -m^2c^2$, which implies the relativistic energy-momentum relation (1.66),

$$E^2 = c^2\vec{p}^{\,2} + m^2c^4 . \tag{2.176}$$

We first leave the electron momentum $\vec{p}$ arbitrary and later transform into the frame in which the electron is initially at rest. The $\mu = 0$ component of (2.174) gives

$$E + \hbar\omega = E' + \hbar\omega' , \tag{2.177}$$

while its spatial components give

$$c\vec{p} + \hbar\omega\hat{e} = c\vec{p}\,' + \hbar\omega'\hat{e}' \; . \tag{2.178}$$

Squaring (2.178), using the relativistic energy-momentum relation (2.176) and eliminating $\vec{p}\,'$ through (2.178) yields

$$E'^2 = E^2 + 2\hbar c\vec{p}\cdot(\omega\hat{e} - \omega'\hat{e}') + \hbar^2\left(\omega\hat{e} - \omega'\hat{e}'\right)^2 \; . \tag{2.179}$$

Next, we use (2.177) to eliminate E' and find after brief rearranging

$$E\left(\omega - \omega'\right) = \hbar\omega\omega'(1 - \cos\theta) + c\vec{p}\cdot(\omega\hat{e} - \omega'\hat{e}') \; , \tag{2.180}$$

where the scattering angle θ of the photon was introduced by $\cos\theta = \hat{e}\cdot\hat{e}'$.

?

Convince yourself of the result (2.180) by your own calculation.

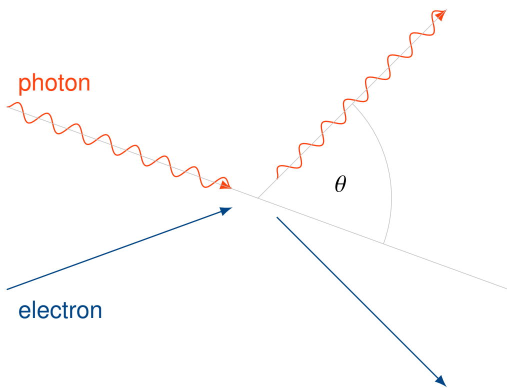

Figure 2.12 Sketch of the kinematics of a Compton-scattering event. The total incoming four-momentum is conserved.

Let us now transform into the rest frame of the electron before the scattering event. There, we can set $\vec{p} = 0$ and $E = mc^2$ in (2.180). The remaining equation is quickly solved for the frequency of the photon after scattering,

$$\frac{\omega'}{\omega} = \frac{1}{1 + \varepsilon(1 - \cos\theta)} \; , \tag{2.181}$$

where $\varepsilon = \hbar\omega/E_0$ is the energy ratio between the photon energy and the electron's rest-energy. Averaging this last result over angles, taking the unpolarised Thomson cross section (2.14) into account, we find the mean relative frequency or energy change per photon,

$$\begin{aligned} \frac{\langle\Delta E_\gamma\rangle}{E_\gamma} = \frac{\langle\omega'\rangle}{\omega} - 1 &= \frac{1}{\sigma_\mathrm{T}}\frac{r_\mathrm{e}^2}{2}\int\frac{\left(1 + \cos^2\theta\right)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi}{1 + \varepsilon(1 - \cos\theta)} - 1 \\ &= \frac{\pi r_\mathrm{e}^2}{\sigma_\mathrm{T}}\int_{-1}^{1}\frac{\left(1 + \mu^2\right)\mathrm{d}\mu}{1 + \varepsilon(1 - \mu)} - 1 \\ &= \frac{\pi r_\mathrm{e}^2}{\sigma_\mathrm{T}}\frac{\ln(1 + 2\varepsilon)\left(2\varepsilon^2 + 2\varepsilon + 1\right) - 2\varepsilon(1 + \varepsilon)}{\varepsilon^3} - 1 \; . \end{aligned} \tag{2.182}$$

Notice that no approximation has so far been made in the rest frame of the electron prior to scattering. Now, we introduce the often appropriate limiting case of photons whose energy is much below the rest energy of the electron, $\varepsilon \ll 1$. Then, by Taylor-expanding the ε-dependent first term in (2.182) to second order, we find the energy change per photon per Compton-scattering event

$$\frac{\langle \Delta E_\gamma \rangle}{E_\gamma} \approx \frac{8\pi r_\mathrm{e}^2}{3\sigma_\mathrm{T}}(1-\varepsilon) - 1 = -\varepsilon = -\frac{\hbar\omega}{mc^2} \ . \tag{2.183}$$

This is our first important result: The relative energy loss of a photon scattering off an electron is given by the ratio of the photon energy and the rest energy of the electron.

?

Show that the result (2.182) is correct and that second-order Taylor approximation in ε leads to (2.183).

2.6.2 Net energy transfer

We now have two competing effects. An electron moving through a sea of radiation is accelerated by the Lorentz force of the electromagnetic radiation field, hence it radiates and transfers the power given by (2.163) to the radiation field. At the same time, photons transfer part of their energy through Compton collisions back to the electrons. For comparing both effects, we first transform our previous result (2.163) from an energy loss per electron per unit time to an energy increase per photon per unit time.

When we studied the energy transfer from a moving charge to an isotropic radiation field, we saw that the power transferred from the electron is proportional to the energy density U of the radiation field. Let now U_ω be the specific energy density of the radiation field contributed by photons with frequency ω. We must then satisfy the normalisation condition

$$U = \int U_\omega \mathrm{d}\omega \ . \tag{2.184}$$

According to (2.163), a single electron increases the energy in such photons by the amount

$$\frac{\mathrm{d}E_\omega^+}{\mathrm{d}t} = \frac{4}{3}\beta^2\gamma^2 c U_\omega \sigma_\mathrm{T} \tag{2.185}$$

per unit time. Let the number density of electrons with velocity β be $n_\mathrm{e}(\beta)$, and the total number density of all electrons be

$$n_\mathrm{e} = \int_0^\infty \mathrm{d}\beta\, n_\mathrm{e}(\beta) \ . \tag{2.186}$$

Then, the electrons contained in a unit of volume increase the energy density in photons with frequency ω by the amount

$$\frac{\mathrm{d}U_\omega^+}{\mathrm{d}t} = \frac{4}{3} n_\mathrm{e}(\beta)\beta^2\gamma^2 c U_\omega \sigma_\mathrm{T} \tag{2.187}$$

per unit time, irrespective of the photon frequency. Since the spatial number density of photons of frequency ω is

$$n_\gamma(\omega) = \frac{U_\omega}{\hbar\omega} \ , \tag{2.188}$$

the sought energy gained per photon per unit time from the electrons with velocity β is

$$\frac{\mathrm{d}E_\gamma^+}{\mathrm{d}t} = \frac{\mathrm{d}U_\omega^+}{\mathrm{d}t} n_\gamma^{-1}(\omega) = \frac{4}{3}\beta^2\gamma^2 n_\mathrm{e}(\beta) c\hbar\omega\sigma_\mathrm{T} \ . \tag{2.189}$$

To find the total energy gain of the photons due to the complete electron population with number density n_e, we need to integrate over the velocity β, see (2.186). Define the average of $\beta^2\gamma^2$ by

$$\left\langle\beta^2\gamma^2\right\rangle = n_\mathrm{e}^{-1}\int_0^\infty \mathrm{d}\beta\,\beta^2\gamma^2 n_\mathrm{e}(\beta)\ , \tag{2.190}$$

then the energy gain per photon and unit time due to electrons of all velocities is

$$\frac{\mathrm{d}E_\gamma^+}{\mathrm{d}t} = \frac{4}{3}\left\langle\beta^2\gamma^2\right\rangle n_\mathrm{e}c\hbar\omega\sigma_\mathrm{T}\ . \tag{2.191}$$

Similarly, the number of Compton collisions that a photon experiences with electrons of total number density n_e is $cn_\mathrm{e}\sigma_\mathrm{T}$. According to (2.183), the energy change per photon per unit time is

$$\frac{\mathrm{d}E_\gamma^-}{\mathrm{d}t} = -cn_\mathrm{e}\sigma_\mathrm{T}\frac{(\hbar\omega)^2}{mc^2}\ . \tag{2.192}$$

Now we can compare the energy gained per photon per unit time, expressed by (2.191), with the energy loss (2.192) per photon per unit time. The total energy change per photon per unit time is the sum of gain and loss,

$$\frac{\mathrm{d}E_\gamma}{\mathrm{d}t} = \frac{\mathrm{d}E_\gamma^+}{\mathrm{d}t} + \frac{\mathrm{d}E_\gamma^-}{\mathrm{d}t} = cn_\mathrm{e}\sigma_\mathrm{T}\hbar\omega\left(\frac{4}{3}\left\langle\beta^2\gamma^2\right\rangle - \frac{\hbar\omega}{mc^2}\right)\ . \tag{2.193}$$

2.6.3 The Kompaneets equation

An illustrative combination of the Fokker-Planck approach and Compton scattering leads to an evolution equation for the phase-space density of photons passing through a hot electron gas. This is most useful in the context of the Cosmic Microwave Background (CMB). The CMB decouples from the quickly recombining cosmic plasma when its temperature falls to $\approx$ 3000 K, corresponding to a thermal energy of $\approx$ 0.3 eV. After that, the CMB photons are redshifted by a factor of $\approx 100\ldots1000$ before they propagate through plasma inside galaxies or galaxy clusters. They have thus typical thermal energies in the meV range or further below. The electron energies even in relatively cool plasmas are typically higher by factors $\gtrsim 10^6$, but still well non-relativistic. In such circumstances, it is appropriate to study Compton scattering under the approximations

$$\frac{\hbar\omega}{c} \ll p_\mathrm{e} \ll mc\ , \tag{2.201}$$

where p_e is the electron momentum.

Let us return with these approximations to the exact equation (2.180) for the frequency change of the scattered photon and stay in the laboratory frame, thus leave $\vec{p}_\mathrm{e} \neq 0$. Due to our approximations, we can then neglect the first term on the right-hand side of (2.180) and write

$$\omega - \omega' \approx \frac{c\vec{p}_\mathrm{e}\cdot(\omega\hat{e} - \omega'\hat{e}')}{mc^2}\ . \tag{2.202}$$

Example: Thermal equlibrium between electrons and photons

For a specific example, suppose now that the electrons have a thermal velocity distribution with a temperature T_e such that $kT_\mathrm{e} \ll mc^2$. The electrons are then non-relativistic, allowing us to set $\gamma \approx 1$. By the equipartition theorem, for systems in thermal equilibrium, their mean-squared velocity must be

$$\left\langle \beta^2 \right\rangle = \frac{3kT_\mathrm{e}}{mc^2} \approx \left\langle \beta^2\gamma^2 \right\rangle \;. \tag{2.194}$$

For averaging the energy gain (2.191) over all photon frequencies, we need to adopt a photon spectrum and calculate the mean energy $\langle \hbar\omega \rangle$ as well as the mean squared energy $\langle (\hbar\omega)^2 \rangle$. Suppose that the photons have a Planck spectrum with temperature T_γ. In terms of the dimension-less energy parameter

$$x := \frac{\hbar\omega}{kT_\gamma} \;, \tag{2.195}$$

the number of photon states in an infinitesimally thin spherical shell with radius x and width $\mathrm{d}x$ is

$$n_x(T_\gamma)\mathrm{d}x = \frac{1}{\pi^2}\left(\frac{kT_\gamma}{\hbar c}\right)^3 \frac{x^2\mathrm{d}x}{\exp(x) - 1} \tag{2.196}$$

according to the Bose-Einstein occupation number in (2.392). By means of the integral

$$\int_0^\infty \frac{x^n\mathrm{d}x}{\exp(x) - 1} = n!\zeta(n + 1) \;, \tag{2.197}$$

the moments of the photon-energy distribution can be calculated to be

$$\langle \hbar\omega \rangle = kT_\gamma \frac{3\zeta(4)}{\zeta(3)} \;, \quad \langle (\hbar\omega)^2 \rangle = (kT_\gamma)^2 \frac{12\zeta(5)}{\zeta(3)} \;. \tag{2.198}$$

When inserted into (2.193) together with the mean-squared velocity (2.194) of the electrons, they give the mean energy gain per photon per unit time due to thermal electrons,

$$\left\langle \frac{\mathrm{d}E_\gamma}{\mathrm{d}t} \right\rangle = \frac{12\zeta(4)}{\zeta(3)} cn_\mathrm{e}\sigma_\mathrm{T} \frac{(kT_\gamma)(kT_\mathrm{e})}{mc^2} \left(1 - \frac{\zeta(5)T_\gamma}{\zeta(4)T_\mathrm{e}}\right) \;. \tag{2.199}$$

This is a highly intriguing result: The energy transfer between thermal populations of electrons and photons should vanish if the temperature of the electrons was slightly *higher* than that of the photons,

$$\frac{T_\mathrm{e}}{T_\gamma} = \frac{\zeta(5)}{\zeta(4)} \;, \tag{2.200}$$

even if the ratio between the temperatures is near unity? This could imply one of two conclusions: Either, finite energy transfer from the photons to the electrons would remain in thermal equilibrium between the two species, *defined* to occur at equal temperatures, or the net energy transfer would cease if the two species were slightly *out of* thermal equilibrium? ◀

Example: Thermal equlibrium between electrons and photons (continued)

Needless to say, a perpetuum mobile could be constructed if either one of these conclusions would be correct, but a perpetuum mobile is forbidden by the second law of thermodynamics. Therefore, the result (2.199) cannot be quite right. The error sneaked in when, in (2.196), we assumed a Bose-Einstein distribution for the photons with vanishing chemical potential, $\mu = 0$. The conclusion from (2.199), combined with the second law of thermodynamics, is therefore much more interesting: If a photon and an electron population coexist in thermal equilibrium, the photons must acquire a finite chemical potential. Then, they cannot maintain their Planck spectrum any longer, but must obtain a spectrum that is slightly deformed by the finite chemical potential. ◀

The energy change of the photon will thus also be small, and we can proceed to approximate

$$\vec{p}_e \cdot (\omega\hat{e} - \omega'\hat{e}') \approx \omega\vec{p}_e \cdot (\hat{e} - \hat{e}') = \omega p_e \left|\hat{e} - \hat{e}'\right| \cos\theta \,, \tag{2.203}$$

where we have introduced the angle θ between the electron momentum $\vec{p}_e$ and the vector $(\hat{e} - \hat{e}')$. Since the modulus of the difference vector $(\hat{e} - \hat{e}')$ is

$$\left|\hat{e} - \hat{e}'\right| = \sqrt{2 - 2\cos\theta} \,, \tag{2.204}$$

we can write (2.202) as

$$\delta\omega \approx -\frac{\omega p_e}{mc} \cos\theta \sqrt{2 - 2\cos\theta} \,. \tag{2.205}$$

This is a typical case suggesting a treatment with the Fokker-Planck approach. The change of the phase-space density $f(\omega)$ of the photons with time is then described by the radial Fokker-Planck equation (1.175)

$$\frac{\partial f}{\partial t} + \frac{1}{p^2}\frac{\partial\left(j_p p^2\right)}{\partial p} = 0 \,, \tag{2.206}$$

where p is the photon momentum. The current density of the radial photon momentum is given by (1.177),

$$j_p = D_2 f \frac{\partial}{\partial p}\left(\ln f - \ln \bar{f}\right) \,. \tag{2.207}$$

To be specific, the distributions f and $\bar{f}$ are the actual and the equilibrium phase-space distributions of the photons. In thermal equilibrium with the electrons, the photons would attain a Bose-Einstein distribution with the appropriate chemical potential and the temperature of the electrons, T_e, which is many orders of magnitude larger than the actual photon temperature. For this reason, the term involving $\bar{f}$ in (2.207) can be neglected altogether in our application, allowing us to approximate simply

$$j_p \approx D_2 \frac{\partial f}{\partial p} \,. \tag{2.208}$$

This leaves the Fokker-Planck equation (2.206) in the simple form

$$\frac{\partial f}{\partial t} + \frac{1}{p^2}\frac{\partial}{\partial p}\left(D_2 p^2 \frac{\partial f}{\partial p}\right) = 0 \,. \tag{2.209}$$

Next, we need to work out the diffusion coefficient D_2. As we have emphasised in Sect. 1.4.4, its physical meaning is one half of the mean-squared momentum change per unit time of the population of scattered particles, i.e. of the photons in the present case. From the frequency change per scattering (2.205), we find the mean-squared momentum change

$$\begin{aligned} D_2 &= \frac{1}{2}\left\langle \delta p^2 \right\rangle = \frac{1}{2}\frac{\hbar^2}{c^2}\left\langle \delta\omega^2 \right\rangle \\ &= \left(\frac{\hbar\omega}{mc^2}\right)^2 \left\langle p_\mathrm{e}^2 \cos^2\theta \right\rangle n_\mathrm{e} c \int \mathrm{d}\Omega \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(1-\cos\theta) \,. \end{aligned} \tag{2.210}$$

By the equipartition theorem, an electron population in thermal equilibrium must have the mean-squared momentum

$$\left\langle p_\mathrm{e}^2 \right\rangle = 2m\,\frac{3}{2}kT_\mathrm{e} = 3mkT_\mathrm{e} \,, \tag{2.211}$$

while the mean-squared $\cos\theta$ gives a factor of $1/3$. For the differential cross section, we use the unpolarised Thomson cross section (2.14),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{r_\mathrm{e}^2}{2}\left(1+\cos^2\theta\right) \,. \tag{2.212}$$

The solid-angle integral in (2.210) then simply gives the total Thomson cross section

$$\sigma_\mathrm{T} = \frac{8\pi}{3} r_\mathrm{e}^2 \,. \tag{2.213}$$

Taking all factors together, we obtain

$$D_2 = \left(\frac{\hbar\omega}{mc^2}\right)^2 mcn_\mathrm{e}\sigma_\mathrm{T}kT_\mathrm{e} = \frac{p^2}{mc} n_\mathrm{e}\sigma_\mathrm{T}kT_\mathrm{e} \,. \tag{2.214}$$

Putting this result back into the Fokker-Planck equation (2.209), we find

$$\frac{\partial f}{\partial t} + cn_\mathrm{e}\sigma_\mathrm{T}\frac{kT_\mathrm{e}}{mc^2}\frac{1}{p^2}\frac{\partial}{\partial p}\left(p^4\frac{\partial f}{\partial p}\right) = 0 \,. \tag{2.215}$$

Let us finally replace the time by the so-called Compton parameter y, defined by

$$\mathrm{d}y = \frac{kT_\mathrm{e}}{mc^2}\, cn_\mathrm{e}\sigma_\mathrm{T}\mathrm{d}t \,. \tag{2.216}$$

This has an intuitive physical meaning: The first factor is the relative energy change of a photon with energy kT by Compton scattering; cf. (2.183). The second factor is the probability for a photon experiencing a Compton-scattering event within the time interval $\mathrm{d}t$. Thus, the differential Compton-y parameter quantifies the mean relative energy change of a photon within the time interval $\mathrm{d}t$. It allows us to bring the Fokker-Planck-equation into the form

$$\frac{\partial f}{\partial y} + \frac{1}{p^2}\frac{\partial}{\partial p}\left(p^4\frac{\partial f}{\partial p}\right) = 0 \,. \tag{2.217}$$

This is *not* exactly the so-called Kompaneets equation, which is often derived and used in this context. However, it is the appropriate limit of the Kompaneets equation, which reveals its origin from the much more general approach of Fokker-Planck theory.

Let us now insert the Bose-Einstein distribution for the photons with vanishing chemical potential,

$$f = \frac{1}{e^x - 1} \quad \text{with} \quad x := \frac{cp}{kT} \tag{2.218}$$

into the Kompaneets equation (2.217). Since p appears to fourth order in the numerator as well as the denominator in the second term of (2.217), we can replace p by x directly. Further, we use

$$f' = -f^2 e^x \quad \text{and} \quad f'' = -f e^x (f + 2f') \ . \tag{2.219}$$

After brief rearrangement, this turns the Kompaneets equation into

$$\frac{\partial f}{\partial y} = \frac{x e^x}{(e^x - 1)^2}\left(x\frac{e^x + 1}{e^x - 1} - 4\right) \ . \tag{2.220}$$

?

The complete Kompaneets equation reads

$$\frac{\partial f}{\partial y} + \frac{1}{p^2}\frac{\partial}{\partial p}\left[p^4\left(\frac{\partial f}{\partial p} + f + f^2\right)\right] = 0 \ .$$

Where could the additional terms arise from, and why is (2.217) an appropriate limit for our purposes?

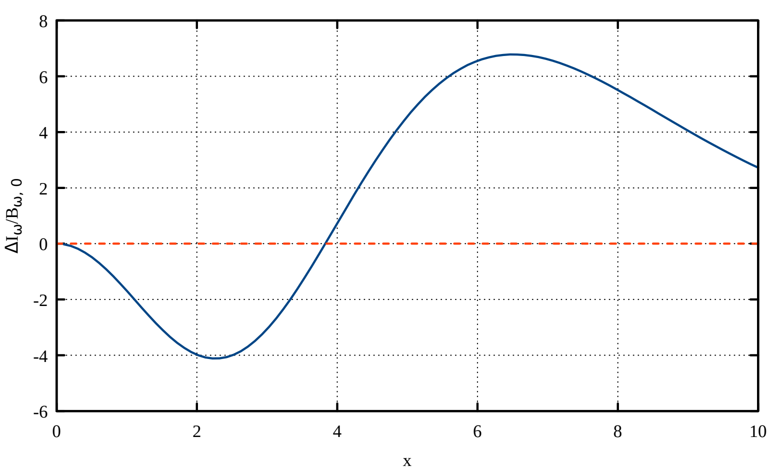

Figure 2.13 The relative change $\Delta I_\omega / B_{\omega,0}$ of the intensity of a black-body spectrum due to Compton scattering is shown as a function of the dimension-less frequency $x = \hbar\omega/(k_B T)$. The intensity is lowered at frequencies below $x = 3.83$ and increased above.

Aiming at astrophysical applications, we are not quite done yet. Notice that (2.220) describes the change of the phase-space density (or the occupation number) of the photons with the Compton-y parameter as they propagate through a plasma. As we shall show below, the intensity is related to the occupation number by (2.396). To obtain the change of intensity with the Compton-y parameter instead, we need to multiply the Kompaneets equation (2.220) by a factor $B_{\omega,0}x^3$, with amplitude $B_{\omega,0}$ of the Planck spectrum defined in (2.401). Thus, we find after integration over y

$$\Delta I_\omega = B_{\omega,0}\frac{x^4 e^x}{(e^x - 1)^2}\left(x\frac{e^x + 1}{e^x - 1} - 4\right) y =: B_{\omega,0}\, g(x)\, y \ . \tag{2.221}$$

This intensity change (Figure 2.13) has an intuitive origin: By Compton scattering, photons are neither created nor destroyed, but only re-distributed in frequency. Based on our initial assumption (2.201), we have studied the effect of high-energy electrons scattering low-energy photons. By the inverse Compton effect, the electrons scatter way more photons from low to high energy rather than the other way. The net effect is thus a depletion of photons relative to the Planck spectrum at low frequencies, and an enhancement at high frequencies. The division between low and high frequencies is set by the root of the function $g(x)$ defined in (2.221), which is numerically found to be at $x_0 = 3.83$. For the Planck spectrum of the CMB, we shall see in (2.414) that the frequency characteristic for its temperature is

$$\nu_{\rm CMB} = \frac{k_{\rm B}T_{\rm CMB}}{h} = 56.8\,\text{GHz}\,, \tag{2.222}$$

which allows to convert x_0 to the frequency

$$\nu_0 = x_0\,\frac{k_{\rm B}T_{\rm CMB}}{h} = 217.5\,\text{GHz}\,. \tag{2.223}$$

Any hot plasma between us and the CMB will therefore reduce the specific CMB intensity below 217.5 GHz, and enhance it above.

Perhaps the most prominent example of huge bodies of hot plasma on the way between the CMB and us are galaxy clusters whose plasma has temperatures of $1\,\text{keV} \lesssim k_{\rm B}T \lesssim 10\,\text{keV}$ and radii of order $R \approx 1\,\text{Mpc} \approx 3.1 \cdot 10^{24}\,\text{cm}$. Their electron number densities are typically $n_{\rm e} \approx 10^{-2}\,\text{cm}^{-3}$. A crude estimate for their Compton-y parameter is

$$y \approx \frac{k_{\rm B}T}{m_{\rm e}c^2}\sigma_{\rm T}n_{\rm e}R \approx 10^{-4}\,. \tag{2.224}$$

Galaxy clusters thus have a very specific spectral signature against the CMB: They cast shadows on the CMB below 217.5 GHz and appear as sources above. The amplitude of the shadows and the sources are of order a milli-Kelvin. This thermal Sunyaev-Zel'dovich effect has turned into an important means for discovering and probing galaxy clusters.

Problems

1. Carry out the steps leading from (2.178) to (2.180).
2. Electrons passing through a plasma lose energy also by Coulomb scattering, i.e. by their interaction with ions through the Coulomb force. A detailed treatment of the Coulomb scattering process shows that the relative energy loss in a single Coulomb-scattering event is

$$\frac{\Delta E}{E} = 4\,\frac{m_{\rm e}}{m_{\rm i}}\,, \tag{2.225}$$

irrespective of the impact parameter.

(a) Derive the ratio between the remaining energy of an electron and its initial energy after n Coulomb collisions.

(b) Approximating $m_e \ll m_i$, how many collisions are needed for the electron to lose half its initial energy?

3. The differential cross section for photons with energy $\hbar\omega$ that are scattered off free electrons is given by the Klein-Nishina formula

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2} F^2(\omega, \theta) \left[F(\omega, \theta) + \frac{1}{F(\omega, \theta)} - 1 + \cos^2\theta \right] , \qquad (2.226)$$

where r_e is the classical electron radius and

$$F(\omega, \theta) = \left[1 + \frac{\hbar\omega}{m_e c^2}(1 - \cos\theta) \right]^{-1} . \qquad (2.227)$$

(a) What is the ratio $\hbar\omega/m_e c^2$ for visible light? How does the Klein-Nishina formula simplify in this case? Is the solution familiar to you?

(b) Assume that an electron is hit by a γ photon with $\hbar\omega = m_e c^2$. Calculate the total cross section

$$\sigma_{KN} = \int d\Omega \frac{d\sigma}{d\Omega} \qquad (2.228)$$

and compare it to the classical Thomson cross section σ_T.

4. Consider a photon with frequency ω scattered by a resting electron under the angle θ. By the scattering process, its frequency changes to $\omega' < \omega$. One can transform into the barycentre system, defined by $\vec{p}_{tot} = 0$ before and after the scattering, by applying a proper Lorentz boost

$$(\Lambda^\mu{}_\nu) = \begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix} \qquad (2.229)$$

to the four-momentum $(p^\mu) = (E/c, \vec{p})^T$, assuming that the incoming photon moves along the negative z-direction.

(a) Calculate the energies and momenta of both the electron and the photon in the barycentre system as a function of β.

(b) Determine the velocity β as a function of ω and the electron mass m_e.

(c) Express the scattering angle θ^* in the barycentre system as a function of the scattering angle θ in the rest frame of the electron, ω and m_e.

2.7 Radiative Quantum Transitions

This section deals with the interaction of electromagnetic radiation with quantum systems such as atoms or ions. We first derive the interaction

Hamiltonian (2.262) by a semi-classical approach treating the electromagnetic field as a classical rather than a quantum field. Next, we relate the amplitude of the interaction Hamiltonian to the intensity of the incoming radiation, enabling us to express the quantum-mechanical transition probability (2.270) by the intensity and the transition matrix element between the initial and the final state. We then introduce the dipole approximation and simplify the transition probability (2.291) accordingly. Cross sections for quantum transitions are defined, and expressions for bound-bound and bound-free transitions are given in (2.291) and (2.291).

2.7.1 Transition probability

Up to this point, we have treated electromagnetic radiation either as composed of classical electromagnetic waves, as in Thomson scattering and our treatment of continuous emission spectra, or as a stream of photons, as in Compton scattering. From both points of view, the particles interacting with the radiation had no internal structure. Effects of radiation on their internal structure, or radiative processes caused by transitions between internal configurations, were neglected so far.

We now proceed to see how electromagnetic radiation can cause transitions between quantum states, e.g. in atoms, but also between bound and free electron states. We begin by recalling a result from time-depedent perturbation theory in quantum mechanics.

Suppose the Hamiltonian $\hat{H}$ of a quantum-mechanical system can be decomposed into a time-independent part $\hat{H}^{(0)}$ and a time-dependent perturbation $\hat{H}^{(1)}(t)$,

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{H}^{(1)}(t) \ . \tag{2.230}$$

Let the time-dependent eigenstates of the unperturbed Hamiltonian $\hat{H}^{(0)}$ with eigenvalue E_n be

$$|n(t)\rangle = |n\rangle e^{-iE_n t/\hbar} \ , \tag{2.231}$$

where the state vector $|n\rangle$ does not depend on time. We expand the eigenstates $|\psi_n(t)\rangle$ of the complete Hamiltonian $\hat{H}(t)$ into eigenstates of the unperturbed Hamiltonian,

$$|\psi_n(t)\rangle = \sum_k c_{nk} |n(t)\rangle \ , \tag{2.232}$$

and demand that they solve Schrödinger's equation,

$$i\hbar |\psi_n(t)\rangle = \left[\hat{H}^{(0)} + \hat{H}^{(1)}(t)\right] |\psi_n(t)\rangle \ . \tag{2.233}$$

?

Carry out all steps leading from Schrödinger's equation (2.233) to the evolution equation (2.235) yourself.

In a first step, this leads to

$$i\hbar \left(\dot{c}_{nk} - c_{nk} \frac{iE_k}{\hbar}\right) |n(t)\rangle = \sum_k c_{nk} \left[E_k + \hat{H}^{(1)}(t)\right] |n(t)\rangle \tag{2.234}$$

since the $|n\rangle(t)$ are eigenstates of the unperturbed Hamiltonian $\hat{H}^{(0)}$. Now, we multiply by $\langle m|$ and use the orthonormality of the unperturbed eigenstates to arrive at

$$\dot{c}_{nm} = -\frac{i}{\hbar} \sum_k c_{nk} \langle m|\hat{H}^{(1)}(t)|k\rangle\, e^{i\omega_{mn} t} \ , \tag{2.235}$$

where

$$\omega_{mn} = \frac{E_m - E_n}{\hbar} \tag{2.236}$$

is the frequency associated with the difference between the energy eigenvalues of the unperturbed states $|n\rangle$ and $|m\rangle$.

The evolution equation (2.235) for the expansion coefficients c_{nm} is exact, but in general difficult to solve. To proceed, we assume that the system is in the eigenstate $|n\rangle$ of the unperturbed Hamiltonian when the perturbation sets in at $t = 0$, thus $c_{nk} = \delta_{nk}$, and that the coefficients c_{nk} with $k \neq n$ remain small even while the perturbation is acting. Then, (2.235) simplifies to

$$\dot{c}_{nm} = -\frac{\mathrm{i}}{\hbar} \langle m|\hat{H}^{(1)}(t)|n\rangle\, \mathrm{e}^{\mathrm{i}\omega_{mn}t} \tag{2.237}$$

and can immediately be integrated once the time dependence of the perturbation Hamiltonian $\hat{H}^{(1)}(t)$ is given.

In our context, perturbations by electromagnetic radiation are most important. We can decompose them into monochromatic waves with frequency ω and thus write the perturbation Hamiltonian as

$$\hat{H}^{(1)}(t) = \hat{V}\, \mathrm{e}^{\mathrm{i}\omega t}\, \theta(t) \tag{2.238}$$

with an operator $\hat{V}$ representing the constant amplitude of the wave. The step function $\theta(t)$ expresses that the perturbation is supposed to begin at $t = 0$. Inserting expression (2.238) into (2.237) and integrating leads us to

$$\begin{aligned} c_{nm} &= -\frac{\mathrm{i}}{\hbar}\langle m|\hat{V}|n\rangle \int_0^t \mathrm{d}t'\, \mathrm{e}^{\mathrm{i}(\omega_{mn}-\omega)t'} \\ &= -\frac{V_{mn}}{\hbar(\omega_{mn}-\omega)}\left[\mathrm{e}^{\mathrm{i}(\omega_{mn}-\omega)t} - 1\right] \end{aligned} \tag{2.239}$$

with the transition-matrix element

$$V_{mn} := \langle m|\hat{V}|n\rangle \tag{2.240}$$

of the amplitude $\hat{V}$ of the perturbation Hamiltonian.

The absolute square of c_{nm} is the transition probability into state $|m\rangle$. Dividing this probability by t gives the transition rate Γ. Using

$$1 - \cos x = 2\sin^2\frac{x}{2}\,, \tag{2.241}$$

we find directly from (2.239) the transition rate

$$\Gamma = \frac{|V_{nm}|^2 t}{\hbar^2}\left[\frac{\sin(\omega_{mn}-\omega)t/2}{(\omega_{mn}-\omega)t/2}\right]^2\,. \tag{2.242}$$

?

Can you confirm the expression (2.242) for the transition rate Γ? How could you prove (2.243)?

If we can furthermore take the limit $t \to \infty$, i.e. if the perturbation acts for a time long compared to the time the system takes for the transition from the state $|n\rangle$ to the state $|m\rangle$, we can use

$$\lim_{a\to\infty} a\left(\frac{\sin ax}{ax}\right)^2 = \pi\delta_{\mathrm{D}}(x) \tag{2.243}$$

with $a = t/2$ to bring the transition rate into the form

$$\Gamma = \frac{2\pi|V_{mn}|^2}{\hbar^2}\delta_{\mathrm{D}}\left(\omega_{mn}-\omega\right)\,. \tag{2.244}$$

2.7.2 Perturbing Hamiltonian

Since we are interested in radiative transitions, we need to know the perturbation Hamiltonian belonging to an incident electromagnetic wave. We have seen in (1.59) that the Lagrange function of a free relativistic particle is

$$L = -mc^2 \sqrt{1-\beta^2} \ . \tag{2.245}$$

If the particle has an electromagnetic charge q, it couples to an electromagnetic field. With the four-potential A^μ of the field and the four-velocity u^μ of the particle, the Lagrange function is extended by a coupling term

$$L = \left(-mc^2 + \frac{q}{c} A_\mu u^\mu\right) \sqrt{1-\beta^2} \ . \tag{2.246}$$

?
Determine the equations of motion from the Lagrange function (2.246). Which force term do you expect?

Since the four-potential and the four-velocity have the components

$$A^\mu = \begin{pmatrix} \Phi \\ \vec{A} \end{pmatrix} , \quad u^\mu = \gamma \begin{pmatrix} c \\ \vec{v} \end{pmatrix} , \tag{2.247}$$

this Lagrange function can be written as

$$L = -mc^2 \sqrt{1-\beta^2} - q\Phi + \frac{q}{c} \vec{A} \cdot \vec{v} \ . \tag{2.248}$$

The momentum conjugate to the velocity $\vec{v}$ is

$$\frac{\partial L}{\partial \vec{v}} = \vec{P} = \gamma m \vec{v} + \frac{q}{c} \vec{A} = \vec{p} + \frac{q}{c} \vec{A} , \tag{2.249}$$

where $\vec{p} = \gamma m \vec{v}$ is the momentum of the free particle. The Legendre transform

$$H = \vec{P} \cdot \vec{v} - L \tag{2.250}$$

then turns the Lagrange- into the Hamilton function of a charged particle in an electromagnetic field,

$$H = \frac{1}{2m} \left(\vec{P} - \frac{q}{c} \vec{A}\right)^2 + q\Phi + mc^2 \ . \tag{2.251}$$

According to the correspondence principle, we shall interpret this Hamilton function as a Hamilton operator. In particular, this implies that $\vec{P}$ will have to be replaced by the momentum operator $\hat{P}$,

$$\vec{P} \to \hat{P} = -\mathrm{i}\hbar \vec{\nabla} \ . \tag{2.252}$$

Remaining in quantum mechanics, avoiding the step into quantum electrodynamics, the electromagnetic field components A^μ will be treated as classical fields rather than field operators. Yet, they depend on spatial coordinates $\vec{x}$. These need to be interpreted as position operators, which do not commute with the momentum operator $\hat{P}$. Thus, we also write the vector potential as an operator $\hat{A}$, understanding that this merely reflects that spatial coordinates x_i in the vector potential need to be replaced by position operators $\hat{x}_i$. Expanding the square in (2.251), we thus need to distinguish between

$$\hat{P} \cdot \hat{A} \quad \text{and} \quad \hat{A} \cdot \hat{P} \ . \tag{2.253}$$

However, we have not employed the gauge freedom of electrodynamics yet. Choosing the Coulomb gauge,

$$\vec{\nabla} \cdot \vec{A} = 0 , \tag{2.254}$$

we can pull the momentum operator $\hat{P}$ past the vector-potential operator $\hat{A}$. In addition to and without conflict with the Coulomb gauge, we can further gauge Φ away, $\hat{\Phi} = 0$, and obtain the Hamilton operator

$$\hat{H} = \frac{\hat{P}^2}{2m} + mc^2 - \frac{e}{mc}\hat{A} \cdot \hat{P} + \frac{e^2}{2mc^2}\hat{A}^2 . \tag{2.255}$$

The first two terms reproduce the Hamiltonian $\hat{H}^{(0)}$ of an unperturbed, free particle, if $\hat{P}$ is interpreted as the momentum operator in absence of the electromagnetic field.

Let us now compare the two final terms in (2.255) containing the vector potential. Their ratio η can be estimated by

$$\eta \approx \frac{e}{2c}\frac{A}{P} , \tag{2.256}$$

with typical values A and P of the vector potential and the momentum. In Coulomb gauge with $\Phi = 0$, the electric field is

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} . \tag{2.257}$$

If we decompose $\vec{A}$ into plane waves and consider a single mode with frequency ω,

$$\vec{E} = -\frac{\mathrm{i}\omega}{c}\vec{A} = -\mathrm{i}k\vec{A} = -\frac{2\pi\mathrm{i}}{\lambda}\vec{A} , \tag{2.258}$$

where we have used the dispersion relation $k = \omega/c$ for electromagnetic waves in vacuum. Thus,

$$A \approx \frac{\lambda E}{2\pi} . \tag{2.259}$$

The momentum of the electron in a hydrogen atom is

$$P \approx \alpha mc , \quad \text{where} \quad \alpha = \frac{e^2}{\hbar c} = \frac{1}{137.04} \tag{2.260}$$

is the fine-structure constant. Combining all terms, we estimate

$$\eta \approx \frac{1}{4\pi\alpha}\frac{\lambda e E}{mc^2} . \tag{2.261}$$

The numerator of the second factor is the work done on the electron by a single wave of the incident electromagnetic field. This is compared to the electron's rest energy! Unless the electromagnetic field is so intense that it can deposit a sizeable fraction of the electron's rest energy on the electron by a single wave, we can safely ignore the term quadratic in $\vec{A}$ in (2.255). Our perturbing Hamiltonian is thus

$$\hat{H}^{(1)}(t) = \frac{e}{mc}\hat{A} \cdot \hat{P} = -\mathrm{i}\frac{\hbar e}{mc}\hat{A} \cdot \vec{\nabla} . \tag{2.262}$$

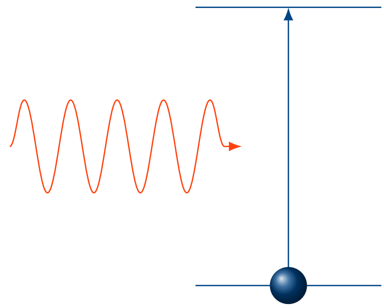

Figure 2.14 Illustration of an incoming electromagnetic wave causing a transition between two quantum states.

2.7.3 Decomposition of the electromagnetic field

Let us now decompose the incident electromagnetic field (cf. Figure 2.14) into plane waves,

$$\vec{A}(\vec{x},t) = A_0\,\hat{e}\,\mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)}\ , \tag{2.263}$$

where A_0 is a scalar, time-independent amplitude and $\hat{e}$ is the polarisation direction. Coulomb gauge immediately implies transversality, $\hat{e}\cdot\vec{k} = 0$. We know that this decomposition into plane waves is possible because the vector potential of electromagnetic waves in vacuum must satisfy the d'Alembert equation $\Box\vec{A} = 0$, what plane waves do if only they obey the dispersion relation $k = \omega/c$.

With (2.244), these plane electromagnetic waves in the perturbing Hamiltonian (2.262) give the transition rate

$$\Gamma = \frac{e^2}{m^2c^2}\,|A_0|^2\,\left|\langle m\,|\mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{x}}\hat{e}\cdot\vec{\nabla}|n\rangle\right|^2\,\delta_{\mathrm{D}}\left(\omega_{mn}-\omega\right)\ . \tag{2.264}$$

We can now relate the absolute square $|A_0|^2$ of the vector-potential amplitude to the intensity of the incoming light. The energy flux density carried by the electromagnetic wave is expressed by its Poynting vector,

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} = \frac{c}{4\pi}\vec{E}^2\hat{e}_k\ , \tag{2.265}$$

where $\hat{e}_k$ is a unit vector pointing into the direction of the wave vector $\vec{k}$. The mean energy flowing past the quantum-mechanical system per unit area and unit time is thus

$$\left\langle\left|\vec{S}\right|\right\rangle = \frac{1}{T}\int_{-T/2}^{T/2}\mathrm{d}t\,\left|\vec{S}\right| = \frac{c}{4\pi T}\int_{-T/2}^{T/2}\mathrm{d}t\,\vec{E}^2\ . \tag{2.266}$$

In the limit of very long times, the time integral can be transformed to a frequency integral by Plancherel's theorem (2.25), which brings (2.266) into the form

$$\left\langle\left|\vec{S}\right|\right\rangle = \frac{c}{4\pi T}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\left|\hat{\vec{E}}\right|^2\ . \tag{2.267}$$

The specific intensity, i.e. the energy per unit area, time and frequency, is thus

$$I_\omega = \frac{c}{8\pi^2 T} \left|\hat{\vec{E}}\right|^2 . \tag{2.268}$$

We can now use (2.258) to continue writing

$$I_\omega = \frac{c}{8\pi^2 T} \frac{\omega^2}{c^2} |A_0|^2 \tag{2.269}$$

and return to the transition rate (2.264) with this result. This gives

$$\Gamma = \frac{8\pi^2 e^2}{m^2 c} \frac{I_\omega T}{\omega^2} \left|\langle m|\mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{x}}\hat{e}\cdot\vec{\nabla}|n\rangle\right|^2 \delta_\mathrm{D}(\omega_{mn} - \omega) \tag{2.270}$$

for the transition rate between the states $|n\rangle$ and $|m\rangle$, given the specific intensity I_ω acting on the system for time T.

2.7.4 Dipole approximation

Before we evaluate the transition matrix element occuring in (2.270), we can apply a further approximation. Expand the phase factor $\exp(\mathrm{i}\vec{k}\cdot\vec{x})$ into a Taylor series,

$$\mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{x}} \approx 1 + \mathrm{i}\vec{k}\cdot\vec{x} - \frac{1}{2}\left(\vec{k}\cdot\vec{x}\right)^2 + \ldots . \tag{2.271}$$

Already the first-order term, $\vec{k}\cdot\vec{r}$, is very much smaller than unity, as the following estimate shows. By the dispersion relation, the wave number k of the electromagnetic wave must be

$$k = \frac{\omega_{mn}}{c} = \frac{E_m - E_n}{\hbar c} , \tag{2.272}$$

while $|\vec{x}| = x$ must be of the order of the Bohr radius a_0,

$$x \approx a_0 = \frac{\hbar^2}{me^2} = 5.2918 \cdot 10^{-9}\,\mathrm{cm} . \tag{2.273}$$

Thus, their product can be estimated to be

$$\vec{k}\cdot\vec{x} \approx kx \approx \frac{\hbar}{me^2 c}(E_m - E_n) = \frac{E_m - E_n}{\alpha mc^2} , \tag{2.274}$$

where we have identified the fine-structure constant α, see (2.260). As long as the energy difference between the transitions is very small compared to the rest-energy of the electron, it is thus very well justified to replace the phase factor by unity. Consider transitions in the hydrogen atom as an example. The ionisation energy of hydrogen is 13.6 eV, while $\alpha mc^2 \approx (511/137)\,\mathrm{keV} \approx 3.7 \cdot 10^3\,\mathrm{eV}$. In this case, $kx \approx 3.7 \cdot 10^{-3}$. We are then left to evaluate the matrix element

$$\langle m|\hat{e}\cdot\vec{\nabla}|n\rangle = \frac{\mathrm{i}}{\hbar}\langle m|\hat{e}\cdot\hat{p}|n\rangle \tag{2.275}$$

Since $|n\rangle$ and $|m\rangle$ are eigenstates of the *unperturbed* Hamiltonian, it is most useful to replace the momentum operator $\hat{p}$ by means of the following commutation relation,

$$\left[\hat{x}, \hat{p}^2\right] = \hat{p}_x[\hat{x}, \hat{p}_x] + [\hat{x}, \hat{p}_x]\hat{p}_x = 2\mathrm{i}\hbar\hat{p}_x . \tag{2.276}$$

Caution Notice that the Bohr radius can be expressed by the classical electron radius (2.12) and the fine-structure constant α as

$$a_0 = \frac{\hbar^2}{me^2} = \frac{e^2}{mc^2}\frac{\hbar^2 c^2}{e^4} = \frac{r_\mathrm{e}}{\alpha^2} .$$

We shall use this relation in (2.300) below. ◀

It allows us to write

$$\left[\hat{x}, \hat{H}\right] = \frac{\mathrm{i}\hbar}{m}\,\hat{p}\,, \tag{2.277}$$

which turns the transition matrix element (2.275) into

$$\langle m\,|\hat{e}\cdot\vec{\nabla}|n\rangle = -\frac{m\omega_{mn}}{\hbar}\langle m\,|\hat{e}\cdot\hat{x}|n\rangle = -\frac{m\omega_{mn}}{e\hbar}\langle m\,|\hat{e}\cdot\hat{d}|n\rangle\,, \tag{2.278}$$

where the dipole operator $\hat{d} = e\hat{x}$ was introduced. For this reason, the approximation $\exp(\mathrm{i}\vec{k}\cdot\vec{x}) \approx 1$ is called the *dipole approximation*.

?

Verify the relations (2.276) and (2.277).

If the dipole matrix element $\langle m\,|\hat{e}\cdot\hat{d}|n\rangle$ vanishes, we need to proceed to the next order in the Taylor expansion of the phase factor, arriving at the level of the so-called quadrupole transitions. The rate (2.270) for dipole transitions has now assumed the form

$$\Gamma = \frac{8\pi^2 I_\omega T}{c\hbar^2}\left|\langle m\,|\hat{e}\cdot\hat{d}|n\rangle\right|^2\,\delta_\mathrm{D}\left(\omega_{mn}-\omega\right)\,. \tag{2.279}$$

Finally, for the frequent case of unpolarised radiation, the mean-squared projection of $\hat{d}$ on $\hat{e}$ gives a factor of $1/3$, and we arrive at

$$\Gamma = \frac{8\pi^2 I_\omega T}{3c\hbar^2}\left|\vec{d}_{mn}\right|^2\,\delta_\mathrm{D}\left(\omega_{mn}-\omega\right)\,, \tag{2.280}$$

where the dipole matrix element $\vec{d}_{mn} = \langle m\,|\hat{d}|n\rangle$ was defined.

2.7.5 Cross sections

We would like to convert the expression (2.280) for the rate of transitions between the states $|n\rangle$ and $|m\rangle$ into an expression for the transition cross section. We shall consider two cases; transitions between two bound states and transitions between a bound and a free state.

Let us begin with transitions between two bound states, which we assume for simplicity to be non-degenerate. Thus, the initial and the final states can be occupied by a single electron each. The two states differ by the discrete energy $E_m - E_n$, which has to be supplied or carried away by a photon with energy $\hbar\omega_{mn} = |E_m - E_n|$. To be specific, we choose to consider the absorption of photons, thus $E_m > E_n$. Of the incoming specific intensity I_ω, only those photons can be absorbed whose frequency precisely equals ω_{mn}. This is expressed by the product $I_\omega\delta_\mathrm{D}(\omega_{mn}-\omega)$ in the transition probability (2.280). Notice that the frequency integral over the Dirac delta function must be dimension-less, so the delta function must have the dimension $[\text{frequency}]^{-1}$. The number of incoming photons at the frequency ω during the time T per area is

$$\frac{I_\omega T}{\hbar\omega}\,. \tag{2.281}$$

Dividing (2.280) by this number gives the desired absorption cross section

$$\sigma_{mn} = \frac{8\pi^2}{3c\hbar}\omega_{mn}\left|\vec{d}_{mn}\right|^2\delta_\mathrm{D}\left(\omega_{mn}-\omega\right)\,, \tag{2.282}$$

with the Dirac delta function expressing that the transition is assumed for now to be needle-sharp in frequency. Conventionally, the dimension-less quantity

$$f_{mn} := \frac{2m\omega_{mn}}{3\hbar e^2}\left|\vec{d}_{mn}\right|^2 \tag{2.283}$$

is called the *oscillator strength* of the transition from the state $|n\rangle$ to the state $|m\rangle$. Identifying it in (2.282) allows us to write the cross section in the simple form

$$\sigma_{mn} = \frac{4\pi^2 e^2}{mc} f_{mn}\,\delta_{\rm D}\left(\omega_{mn}-\omega\right) = 4\pi^2 r_{\rm e} c\, f_{mn}\,\delta_{\rm D}\left(\omega_{mn}-\omega\right)\;, \tag{2.284}$$

where the classical electron radius $r_{\rm e} = 2.81\cdot 10^{-13}$ cm was introduced from (2.12).

?

Is the oscillator strength (2.283) really dimension-less, as claimed?

In realistic situations, as we shall see below, the absorption cross section does not have the needle-sharp delta profile adopted here, but a broader one. If this profile is described by a function $\phi(\omega_{mn}-\omega)$ which is normalised to unity, the cross section reads

$$\sigma_{mn} = \frac{4\pi^2 e^2}{mc} f_{mn}\,\phi\left(\omega_{mn}-\omega\right) = 4\pi^2 r_{\rm e} c\, f_{mn}\,\phi\left(\omega_{mn}-\omega\right)\;. \tag{2.285}$$

As we shall see shortly, there is a characteristic line profile function, called the Voigt profile.

For bound-free transitions, we can proceed in an analogous way as for bound-bound transitions, except that we have to take the number of available free electron states into account. We arrive at the bound-free absorption cross section $\sigma_{\rm bf}$ if we multiply the transition rate (2.270) by the number of final electron states and divide, as before, by the number of photons incoming per unit area per unit time. The number of final electron states in an infinitesimally thin momentum shell in phase space is

$$\frac{4\pi p_{\rm f}^2\,{\rm d}p_{\rm f}}{(2\pi\hbar)^3}V = \frac{p_{\rm f}^2\,{\rm d}p_{\rm f}}{2\pi^2\hbar^3}V = \frac{k_{\rm f}^2\,{\rm d}k_{\rm f}}{2\pi^2}V \tag{2.286}$$

if the shell has the width ${\rm d}p_{\rm f} = \hbar{\rm d}k_{\rm f}$ in the final electron momentum. Energy conservation implies that the energy of an incoming photon, $\hbar\omega$, must come up for the binding energy E_1 of the electron plus the energy of the free electron after ionisation,

$$\hbar\omega = \frac{p_{\rm f}^2}{2m} + E_1 = \frac{\hbar^2 k_{\rm f}^2}{2m} + E_1\;. \tag{2.287}$$

This allows us to relate the width ${\rm d}k_{\rm f}$ of the shell of electron momenta to the width ${\rm d}\omega$ in photon frequency,

$$\hbar{\rm d}\omega = \frac{\hbar^2 k_{\rm f}\,{\rm d}k_{\rm f}}{m} \quad\Rightarrow\quad k_{\rm f}{\rm d}k_{\rm f} = \frac{m\,{\rm d}\omega}{\hbar}\;. \tag{2.288}$$

The number of final electron states can thus be expressed by

$$\frac{k_{\rm f}^2\,{\rm d}k_{\rm f}}{2\pi^2}V = \frac{k_{\rm f} m\,{\rm d}\omega}{2\pi^2\hbar}V\;, \tag{2.289}$$

and the number of photons with frequency within $[\omega, \omega + \mathrm{d}\omega]$ incoming during the time T per unit area is given by

$$\frac{I_\omega T}{\hbar\omega}\mathrm{d}\omega\ . \tag{2.290}$$

Multiplying the transition rate (2.270) with the number of electron states (2.289) and dividing by the number (2.290) of incoming photons gives the cross section

$$\sigma_{\mathrm{bf}} = \frac{4e^2 k_{\mathrm{f}}}{mc}\frac{V}{\omega}\left|\langle f\,|\mathrm{e}^{\mathrm{i}\vec k\cdot\vec x}\hat e\cdot\vec\nabla|b\rangle\right|^2 \tag{2.291}$$

between the bound state $|b\rangle$ and the free state $|f\rangle$, where the transition matrix elements still needs to be worked out.

2.7.6 Photoionisation cross section

To give one specific and simple example for the calculation of a bound-free cross section, we consider the photoionisation of the hydrogen atom from its ground state. In the position representation, the bound and free electron states are given by the wave functions

$$\begin{aligned}\psi_{\mathrm{b}}(\vec x) &= \langle x\,|b\rangle = \left(\pi a_0^3\right)^{-1/2}\mathrm{e}^{-r/a_0}\ ,\\ \psi_{\mathrm{f}}(\vec x) &= \langle x\,|f\rangle = V^{-1/2}\mathrm{e}^{\mathrm{i}\vec k_{\mathrm{f}}\cdot\vec x}\ ,\end{aligned} \tag{2.292}$$

?

Can you confirm that the wave functions (2.292) correctly represent the states of an electron bound in the ground state of a hydrogen atom, and a free electron, respectively? Are they properly normalised?

where the final electron state is assumed to be confined to the volume V.

If the energy difference between the final and initial electron states is small compared to the rest-energy of the electron, i.e. as long as the electron remains non-relativistic, we can evaluate the transition matrix element in dipole approximation. We thus set $\exp(\mathrm{i}\vec k\cdot\vec x) \approx 1$ in (2.291) and use the Hermitian property of the momentum operator to exchange the final and the initial states,

$$\left|\langle f\,|\hat e\cdot\vec\nabla|b\rangle\right|^2 = \left|\langle b\,|\hat e\cdot\vec\nabla|f\rangle^*\right|^2 = \left|\langle b\,|\hat e\cdot\vec\nabla|f\rangle\right|^2\ . \tag{2.293}$$

Inserting the initial and final wave functions, the matrix element is now easily evaluated,

$$\begin{aligned}\langle b\,|\hat e\cdot\vec\nabla|f\rangle &= \left(\pi a_0^3 V\right)^{-1/2}\int\mathrm{d}^3x\,\mathrm{e}^{-r/a_0}\hat e\cdot\vec\nabla\mathrm{e}^{\mathrm{i}\vec k_{\mathrm{f}}\cdot\vec x}\\ &= \left(\pi a_0^3 V\right)^{-1/2}\mathrm{i}\hat e\cdot\vec k_{\mathrm{f}}\int\mathrm{d}^3x\,\mathrm{e}^{-r/a_0+\mathrm{i}\vec k_{\mathrm{f}}\cdot\vec x}\ .\end{aligned} \tag{2.294}$$

The remaining integral is quickly worked out in polar coordinates,

$$\begin{aligned}\int\mathrm{d}^3x\,\mathrm{e}^{-r/a_0+\mathrm{i}\vec k_{\mathrm{f}}\cdot\vec x} &= 2\pi\int_0^\infty r^2\mathrm{d}r\,\mathrm{e}^{-r/a_0}\int_{-1}^1\mathrm{d}\cos\theta\,\mathrm{e}^{\mathrm{i}k_{\mathrm{f}}r\cos\theta}\\ &= 4\pi\int_0^\infty r^2\mathrm{d}r\,\mathrm{e}^{-r/a_0}\frac{\sin(k_{\mathrm{f}}r)}{k_{\mathrm{f}}r}\\ &= \frac{8\pi a_0^3}{\left(1+k_{\mathrm{f}}^2a_0^2\right)^2} \approx \frac{8\pi}{k_{\mathrm{f}}^4 a_0}\ ,\end{aligned} \tag{2.295}$$

where the final approximation is allowed if the energy of the final state is much larger than that of the initial state.

Putting the last results back into the bound-free cross section (2.291), we obtain

$$\sigma_{\rm bf} = \frac{256\pi}{3}\frac{e^2}{mc\omega}\frac{1}{(a_0k_{\rm f})^5} = \frac{256\pi}{3}\frac{\alpha\hbar}{m\omega}\frac{1}{(a_0k_{\rm f})^5}\ , \tag{2.296}$$

where we have averaged over all polarisation directions to replace

$$\left(\hat{e}\cdot\vec{k}_{\rm f}\right)^2 = \frac{1}{3}k_{\rm f}^2\ . \tag{2.297}$$

?

Can we really integrate the radius to infinity in (2.295)? What is the crucial approximation behind doing so? Carry out the final radial integral in (2.295) yourself.

Our previous approximation that the energy of the final state largely exceeds that of the initial state allows us to ignore the binding energy E_1 in (2.287) and to substitute

$$k_{\rm f} = \sqrt{\frac{2m\omega}{\hbar}} \tag{2.298}$$

and bring the photoionisation cross section into the form

$$\sigma_{\rm bf} = \frac{64\pi}{3\sqrt{2}}\frac{\alpha}{a_0^5}\left(\frac{\hbar}{m\omega}\right)^{7/2}\ . \tag{2.299}$$

Rearranging the constants, inserting the Bohr radius (2.273) in the form

$$a_0 = \frac{r_{\rm e}}{\alpha^2} \tag{2.300}$$

with the classical electron radius (2.12) as well as the Rydberg energy

$${\rm Ry} = \frac{me^4}{2\hbar^2} = \frac{\alpha^2mc^2}{2} = 13.6\,{\rm eV}\ , \tag{2.301}$$

we can bring the expression for the bound-free cross section into the more intuitive form

$$\sigma_{\rm bf} = \left(\frac{4}{\alpha}\right)^3\sigma_{\rm T}\left(\frac{\rm Ry}{\hbar\omega}\right)^{7/2} = 1.09\cdot10^{-16}\,{\rm cm}^2\left(\frac{\rm Ry}{\hbar\omega}\right)^{7/2} \tag{2.302}$$

containing the Thomson cross section (2.15). It should be kept in mind, however, that this equation is only valid for photon energies much larger than the Rydberg energy, $\hbar\omega \gg {\rm Ry}$.

Problems

1. The cross section for a transition between an initial state $|n\rangle$ and a final state $|m\rangle$ was derived as

$$\sigma_{mn} = \frac{4\pi}{3c\hbar}\omega_{mn}\left|\vec{d}_{mn}\right|^2\delta_{\rm D}\left(\omega_{mn}-\omega\right)\ , \tag{2.303}$$

where $\vec{d}_{mn} = \langle m|e\hat{x}|n\rangle$ is the dipole matrix element and $\omega_{mn} = (E_m - E_n)/\hbar$ is the frequency corresponding to the energy difference between the states $|m\rangle$ and $|n\rangle$. The delta distribution assures that only those photons contribute to the cross section that have the correct frequency.

(a) Consider the one-dimensional harmonic oscillator with energy levels $E_n = \hbar\omega(n + 1/2)$ and corresponding wave functions

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \exp\left(-\frac{m\omega}{2\hbar}x^2\right) \quad (2.304)$$

with the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \ . \quad (2.305)$$

Calculate the cross section σ_{10} for the transition from the ground state ($n = 0$) to the first excited state ($n = 1$). *Hint*: It may be helpful to use

$$\int_{-\infty}^{\infty} dx\, x^2 e^{-\alpha x^2} = -\int_{-\infty}^{\infty} dx \frac{\partial}{\partial\alpha} e^{-\alpha x^2} \ . \quad (2.306)$$

(b) Consider now an infinitely deep potential well of length L with energy levels

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2} \quad (2.307)$$

with $n \in \mathbb{N}$ and wave functions

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \cos\left(\frac{n\pi}{L}x\right) & \text{if } n \text{ is odd} \\ \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right) & \text{if } n \text{ is even} \end{cases} \ , \quad (2.308)$$

with $x \in [-L/2, L/2]$. What is the cross section σ_{21} for the transition from the ground state ($n = 1$) to the first excited state ($n = 2$)? Compare the factor in front of the delta distribution with that for the harmonic oscillator.

2.8 Shapes of Spectral Lines

In this section, three different statements on spectral lines are derived and applied. First, it is shown that spontaneous transitions between quantum states broaden spectral lines emitted by electromagnetic transitions between these states from the needle-sharp profile expected for ideally sharp transitions to a Lorentz profile whose width is determined by the spontaneous transition rate. The first main result is the Lorentzian profile function (2.318). Collisions between emitting quantum systems are shown to have the same effect, with the spontaneous transition rate replaced by the collision rate. Second, the Doppler broadening by the motion of the emitting quantum systems leads to the Gaussian line profile (2.330) if the motion is thermal. Third, the combined effects of spontaneous or collisional transitions and Doppler broadening are shown to create the Voigt line profile (2.337). This combined line profile is then used to determine how the equivalent widths of spectral lines change with the number of absorbers, leading to the curve-of-growth described by (2.352).

2.8.1 Natural line width

Consider now two states of a quantum-mechanical system, for simplicity called $|m\rangle$ and $|n\rangle$, which are separated by the energy difference $E_n - E_m > 0$. If the system is in the upper state $|n\rangle$, it has a finite probability to decay spontaneously to the lower state $|m\rangle$. The state $|n\rangle$ thus has a finite lifetime, which causes an uncertainty in its energy E_n. The transition energy between the two states $|m\rangle$ and $|n\rangle$ will thus be distributed around its precise value $E_n - E_m$. We shall now work out the shape of this distribution.

?

Why would an excited state spontaneously decay into a less energetic state?

We begin with the evolution equation (2.235) for the expansion coefficients c_{nm} perturbed state $|\psi(t)\rangle$ in terms of the eigenstates $|k\rangle$ of an unperturbed Hamiltonian,

$$\dot{c}_{nm} = -\frac{\mathrm{i}}{\hbar}\sum_k c_{nk}\langle m|\hat{H}^{(1)}(t)|k\rangle\,\mathrm{e}^{\mathrm{i}\omega_{mn}t} \tag{2.309}$$

and assume a radiative perturbation Hamiltonian $\hat{H}^{(1)}(t)$ with periodic time dependence as in (2.238),

$$\hat{H}^{(1)}(t) = \hat{V}\,\mathrm{e}^{-\mathrm{i}\omega t}\,\theta(t)\,, \tag{2.310}$$

with a time-independent operator $\hat{V}$.

Let us now restrict our attention to a radiative transition between any two states $|n\rangle$ and $|m\rangle$. Their energies E_m and E_n are supposed to satisfy $E_m > E_n$, respectively. Initially, we assume the system to be in the state $|n\rangle$, which could be its ground state, and thus begin the evolution with $c_{nn} = 1$ and $c_{nk} = 0$ for $k \neq n$. Restricting our general result (2.309) to this simplified two-state system, the coefficient c_{nm} evolves in time according to

$$\dot{c}_{nm} = -\frac{\mathrm{i}}{\hbar}\langle m|\hat{V}|n\rangle\,\mathrm{e}^{\mathrm{i}(\omega_{mn}-\omega)t}\,. \tag{2.311}$$

Strictly speaking, c_{nn} is also time dependent, so we would have to solve a system of coupled differential equations for c_{nn} and c_{nm}. In a first step of what could turn into an iterative approach, we now assume that the ground state remains populated as the transitions are going on, hence $c_{nn} = 1$ for all times. This is justified if the transition probability from $|n\rangle$ to $|m\rangle$ is small. Should this be unreasonable in the situation considered, a first solution for $c_{nm}(t)$ can then be inserted into the evolution equation for c_{nn} to determine a correction, and so forth.

Taking, however, $c_{nn} = 1$ for now, we can immediately solve (2.311) by direct integration, enforcing the initial condition $c_{nm} = 0$ at $t = 0$. This gives

$$c_{mn}(t) = -\frac{\langle m|\hat{V}|n\rangle}{\hbar(\omega_{mn}-\omega)}\left[\mathrm{e}^{\mathrm{i}(\omega_{nm}-\omega)t} - 1\right]\,, \tag{2.312}$$

as in (2.239).

However, this result has the problem that it was derived ignoring that the excited state $|m\rangle$ can decay spontaneously. Very much like radioactive decay, we can phenomenologically model such a spontaneous decay by the introducing a contribution

$$\dot{c}_{nm} \to \dot{c}_{nm} - \frac{\Gamma}{2}\,c_{nm} \tag{2.313}$$

into the differential equation (2.311), where the spontaneous decay rate $\Gamma/2$ was inserted with a factor of $1/2$ that will be convenient later. After this ad-hoc modification, c_{nm} is supposed to evolve according to

$$\dot{c}_{nm} = -\frac{\mathrm{i}}{\hbar}\langle m|\hat{V}|n\rangle\, \mathrm{e}^{\mathrm{i}(\omega_{mn}-\omega)t} - \frac{\Gamma}{2}\, c_{nm}\ . \tag{2.314}$$

After bringing this additional term $\Gamma c_{nm}/2$ to the left-hand side and multiplying the equation with $\mathrm{e}^{\Gamma t/2}$, we see that we can write

$$\partial_t\left(c_{nm}\mathrm{e}^{\Gamma t/2}\right) = -\frac{\mathrm{i}}{\hbar}\langle m|\hat{V}|n\rangle\, \mathrm{e}^{[\mathrm{i}(\omega_{mn}-\omega)+\Gamma/2]t}\ . \tag{2.315}$$

Again, we can directly integrate this equation with the same initial condition as before, $c_{nm} = 0$ at $t = 0$. This gives

$$c_{nm}(t) = \frac{\langle m|\hat{V}|n\rangle}{\hbar}\, \frac{\mathrm{e}^{-\Gamma t/2} - \mathrm{e}^{\mathrm{i}(\omega_{mn}-\omega)t}}{(\omega-\omega_{mn}) + \mathrm{i}\Gamma/2}\ . \tag{2.316}$$

After a sufficiently long initial time $t \gg \Gamma^{-1}$, the exponential term in the numerator dies off. Then, the absolute square of c_{nm}, which gives the probability for finding the system in state $|m\rangle$, becomes time-independent and reads

$$|c_{nm}|^2 = \frac{\left|\langle m|\hat{V}|n\rangle\right|^2}{\hbar^2}\, \frac{1}{(\omega-\omega_{mn})^2 + \Gamma^2/4}\ . \tag{2.317}$$

The dependence of the transition probability on frequency is thus described by the Lorentz profile function

$$\phi_\Gamma(\omega-\omega_{12}) = \frac{1}{\pi}\, \frac{\Gamma/2}{(\omega-\omega_{12})^2 + \Gamma^2/4} \tag{2.318}$$

first encountered in (2.153). Recall that the prefactor in (2.318) is chosen such that ϕ_Γ integrates to unity.

2.8.2 Collisional broadening

When a quantum-mechanical system interacts with another in a collision, its phase is randomly changed, or reset. We model this process by assuming that there is a random phase shift $\delta\phi$ in each collision, which we choose to be drawn from the interval $[-\pi, \pi]$. The probability distribution of $\delta\phi$ within this interval is supposed to be flat such that all phase shifts within $[-\pi, \pi]$ are equally likely.

We cannot know the phase shift after a single collision. However, the average phase factor after a single collision must vanish,

$$\left\langle \mathrm{e}^{\mathrm{i}\delta\phi}\right\rangle = 0\ , \tag{2.319}$$

because of the flat distribution of $\delta\phi \in [-\pi, \pi]$. If more than one collision occurs, the mean phase factor will still vanish: The phase shift after N collisions will have a flat distribution over the interval $[-N\pi, N\pi]$, hence the average phase factor will vanish also if an arbitrary number of collisions has occured.

The number of collisions within a given time t can be modelled as a Poisson process. Let Γ_c be the collision rate. Then the expected number of collisions within the time t is $\Gamma_c t$, and the probability for k collisions to occur during that time is given by the Poisson distribution,

$$p_k = \frac{(\Gamma_c t)^k}{k!} e^{-\Gamma_c t} . \tag{2.320}$$

Since the phase factor after time t is zero if *any* collision has happened, the mean phase factor will be

$$\left\langle e^{i\delta\phi} \right\rangle = p_0 = e^{-\Gamma_c t} . \tag{2.321}$$

Since the probability for the system to be in state $|2\rangle$ is given by $|c_{nm}|^2$, this corresponds to modifying the evolution equation for c_{nm} by a term

$$\dot{c}_{nm} = -\frac{\Gamma_c}{2} c_{nm} . \tag{2.322}$$

A comparison to the treatment of the natural line width above, see (2.314) and (2.315), shows that the only change to the previous solution (2.316) for the transition probability $|c_{nm}|^2$ is that the decay rate Γ for spontaneous transitions is replaced by the sum of the spontaneous and the collisional decay rates,

$$\Gamma \to \Gamma + \Gamma_c . \tag{2.323}$$

The shape of the line profile function (2.318) will thus remain unchanged, only its width will be enhanced by an amount determined by the sum of the spontaneous and the collisional decay rates.

2.8.3 Doppler broadening of spectral lines

A further broadening mechanism is caused by the Doppler effect. If the emitting quantum-mechanical systems, e.g. atoms or molecules, move along the line-of-sight with the velocity $v_\|$, we observe the frequency

$$\omega = \omega_0 \left(1 + \frac{v_\|}{c}\right) \tag{2.324}$$

instead of the emitted frequency ω_0. This is the non-relativistic approximation to the Doppler effect, which we can safely use for atoms or molecules moving thermally. In the thermal case, the emitters can further be expected to have a Maxwellian velocity distribution with a width σ_v determined by their temperature. The equipartition theorem demands

$$\frac{m}{2}\sigma_v^2 = \frac{kT}{2} \quad \Rightarrow \quad \sigma_v^2 = \frac{kT}{m} . \tag{2.325}$$

?

If the velocity distribution of the emitting atoms and molecules would follow a power law, what would the line profile look like?

The single velocity component $v_\|$ then has a Gaussian distribution, and the observed line profile is then given by

$$\int_{-\infty}^{\infty} \frac{dv_\|}{\sqrt{2\pi\sigma_v^2}} \delta_D \left[\omega - \omega_0 \left(1 + \frac{v_\|}{c}\right)\right] \exp\left[-\frac{(v_\| - \bar{v})^2}{2\sigma_v^2}\right] , \tag{2.326}$$

where $\bar{v}$ is the mean velocity of the emitting system. Using the identity

$$\delta_D(ax) = \frac{1}{a}\delta_D(x) , \tag{2.327}$$

for the Dirac delta function, the Gaussian line profile proportional to

$$\exp\left[-\frac{1}{2\sigma_v^2}\left(\frac{\omega-\omega_0}{\omega_0}c-\bar{v}\right)^2\right]=\exp\left[-\frac{c^2}{2\sigma_v^2}\left(\frac{\omega-\bar{\omega}}{\omega_0}\right)^2\right] \tag{2.328}$$

emerges. Here, we have defined the centre frequency

$$\bar{\omega}\equiv\omega_0\left(1+\frac{\bar{v}}{c}\right)\,, \tag{2.329}$$

i.e. the average frequency of the line emission, shifted by the Doppler effect due to the mean motion of the emitting or absorbing medium. The line-profile function $\phi(\omega-\omega_0)$ for thermally moving atoms is thus the (normalised) Gaussian

$$\phi(\omega-\omega_0)=\frac{c}{\omega_0\sqrt{2\pi\sigma_v^2}}\exp\left[-\frac{c^2}{2\sigma_v^2}\left(\frac{\omega-\bar{\omega}}{\omega_0}\right)^2\right]\,. \tag{2.330}$$

2.8.4 The Voigt profile

In presence of all three line-broadening effects, i.e. spontaneous, collisional and Doppler broadening, the line profile is a convolution of the Lorentz profile for the line broadened by spontaneous and collisional decays with the Gaussian velocity distribution taking account of the Doppler effect. The combined line profile is thus determined by the integral

$$\frac{1}{\sqrt{2\pi}\sigma_v}\int_{-\infty}^{\infty}\mathrm{d}v_\parallel\,\phi\left[\omega-\omega_{12}\left(1+\frac{v_\parallel}{c}\right)\right]\exp\left(-\frac{v_\parallel^2}{2\sigma_v^2}\right)\,, \tag{2.331}$$

which can be brought into a standard form by a sequence of substitutions. First, we introduce a velocity scale v_0 and a dimension-less velocity q by

$$v_0\equiv\sqrt{2}\sigma_v \quad \text{and} \quad q\equiv\frac{v_\parallel}{\sqrt{2}\sigma_v}=\frac{v_\parallel}{v_0} \tag{2.332}$$

to bring the Gaussian factor in (2.331) into the form

$$\frac{\mathrm{d}v_\parallel}{\sqrt{2\pi}\sigma_v}\exp\left(-\frac{v_\parallel^2}{2\sigma_v^2}\right)=\frac{\mathrm{d}q}{\sqrt{\pi}}\,\mathrm{e}^{-q^2}\,. \tag{2.333}$$

The Lorentz profile (2.318), with the centre frequency shifted by the Doppler effect, reads

$$\phi_\Gamma\left[\omega-\omega_{12}\left(1+\frac{v_\parallel}{c}\right)\right]=\frac{1}{\pi}\frac{\Gamma/2}{\left[\omega-\omega_{12}\left(1+\frac{v_\parallel}{c}\right)\right]^2+\Gamma^2/4}\,. \tag{2.334}$$

The further substitutions of a centred, normalised frequency u and a normalised collision rate a, defined by

$$u\equiv\frac{\omega-\omega_{12}}{\omega_{12}}\frac{c}{v_0} \quad \text{and} \quad a\equiv\frac{\Gamma}{2\omega_{12}}\frac{c}{v_0}\,, \tag{2.335}$$

bring this profile into the form

$$\phi(u)=\frac{c}{\pi\omega_{12}v_0}\frac{a}{(u-q)^2+a^2}\,. \tag{2.336}$$

The result of the convolution (2.331) thus reads

$$\phi(u) = \frac{ac}{\pi\sqrt{\pi}\omega_{12}v_0}\int_{-\infty}^{\infty} \mathrm{d}q\, \frac{\mathrm{e}^{-q^2}}{(u-q)^2 + a^2} \,, \tag{2.337}$$

which is the so-called the Voigt profile (Figure 2.15). Near its centre, this line profile has a Gaussian shape, while its wings retain the Lorentzian shape.

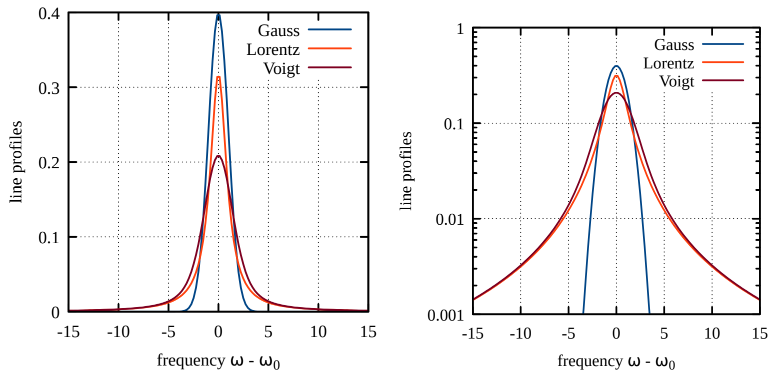

Figure 2.15 The Gauss, the Lorentz and the Voigt profiles are shown for $\sigma = 1$ and $\Gamma = 1$ in arbitrary units. The left and right panels are distinguished only by the linear and logarithmic scaling of the ordinate. The right panel illustrates in particular the broad, Lorentzian wings of the Voigt profile.

2.8.5 Equivalent widths and curves-of-growth

Two concepts have been found useful describing the information contained in observed spectral lines, namely their equivalent width and their curve-of-growth. The equivalent width quantifies the area under a spectral line. If I_0 is the local specific intensity of the spectral continuum, that is the continuum intensity in the vicinity of the line, the equivalent width is defined as

$$W \equiv \int \frac{I_0 - I(\omega)}{I_0} \mathrm{d}\omega \,, \tag{2.338}$$

where $I(\omega)$ is the specific intensity within the line. Thus, the equivalent width of an absorption line is a measure for the total intensity removed from the spectrum. An analogous definition can be given for the equivalent width of emission lines, which then quantifies the total intensity added to the spectrum. The optical depth within the line is given by the number density of absorbers n, the geometrical extent L of the absorbing medium and the frequency-dependent cross section $\sigma(\omega)$,

$$\tau = n\,L\,\sigma(\omega) \,, \tag{2.339}$$

where the specific dependence of $\sigma(\omega)$ on the frequency may be given by (2.279) in dipole approximation. The specific intensity within the line is then lowered compared to the specific continuum intensity I_0 by

$$I(\omega) = I_0\,\mathrm{e}^{-\tau(\omega)} \,, \tag{2.340}$$

?

Why would the specific intensity fall off exponentially with the optical depth, as in (2.340)?

and thus the equivalent width is the integral

$$W = \int \mathrm{d}\omega \left[1 - \mathrm{e}^{-\tau(\omega)}\right] \tag{2.341}$$

across the line.

Since the cross section is proportional to the profile function $\phi(\omega)$, (2.341) can equally well be written as

$$W = \int \mathrm{d}\omega \left[1 - \mathrm{e}^{-C\phi(\omega)}\right] . \tag{2.342}$$

As shown in (2.280), the frequency-independent constant C inserted here is $C = 2\pi n L r_\mathrm{e} c f_{12}$ for a dipole transition between levels 2 and 1 with oscillator strength f_{12}. For low optical depth, $\tau \ll 1$, the exponential function in (2.341) or (2.342) can be replaced by its first-order Taylor expansion. This results in

$$W = \int \mathrm{d}\omega\, n\, L\, \sigma(\omega) = 2\pi n L r_\mathrm{e} c f_{12} \tag{2.343}$$

because the profile function is defined to be normalised such that its integral over frequency ω gives unity. Thus, for low optical depth, we introduce the column density $N = nL$ and have

$$W \propto N , \tag{2.344}$$

i.e. the equivalent width is simply growing linearly with the column density of absorbers along the line-of-sight from the observer.

In the opposite, optically thick case $\tau \gg 1$, the line profile can be approximated by a sudden drop from the continuum level I_0 to zero intensity within a frequency range of width 2Δ, and a sudden rise back to the continuum level. The spectral line is thus simply described as a rectangular stripe cut from the spectrum. By definition of the equivalent width,

$$W \approx 2\Delta \tag{2.345}$$

in this case. We now need to distinguish whether $\tau \approx 1$ is reached only in the core or already in the wings of the spectral-line profile. We first consider the case of an optically thick core, but optically thin wings. If the line is Doppler-broadened, as most lines are, the line profile has a Gaussian core, and we can approximate the optical depth as

$$\tau = NL\sigma(\omega) = 2\pi NL \frac{r_\mathrm{e} c^2 f_{12}}{\sqrt{2\pi}\sigma_v} \exp\left(-\frac{c^2(\omega - \omega_{12})^2}{2\sigma_v^2}\right) \tag{2.346}$$

with a thermal velocity dispersion σ_v given by (2.325). We now determine the half-width $\Delta = \omega - \omega_{12}$ by setting $\tau = 1$ in (2.346) and solving for Δ,

$$\exp\left(-\frac{c^2\Delta^2}{2\sigma_v^2}\right) \stackrel{!}{=} \frac{\sigma_v}{\pi\sqrt{2\pi} N L r_\mathrm{e} c^2 f_{12}} . \tag{2.347}$$

Thus, the width Δ and therefore also the equivalent width scale with N like

$$\Delta \propto \sqrt{\ln N} , \quad W \propto \sqrt{\ln N} . \tag{2.348}$$

i.e. they depend only very weakly on the number N of absorbers.

If, however, $\tau \approx 1$ is reached already in the Lorentzian wings of the line, we adopt the Lorentz profile (2.318) instead of the Voigt profile and further simplify the Lorentz profile by assuming a damping rate Γ small compared to the frequency difference to the line centre, $\Gamma \ll \omega - \omega_{12}$. Then,

$$\phi_\Gamma(\omega) \approx \frac{\Gamma}{2\pi(\omega - \omega_{12})^2} , \tag{2.349}$$

and the optical depth becomes

$$\tau = NL\sigma(\omega) \approx \pi N L r_e c f_{12} \frac{\Gamma}{(\omega - \omega_{12})^2} . \tag{2.350}$$

As above, we find the width Δ from this equation by setting $\tau = 1$ in (2.350) and solving for $\Delta = \omega - \omega_{12}$. This reveals that in this case of very high optical depth, Δ and the equivalent width W scale with the number N of absorbers like

$$\Delta \propto \sqrt{N} , \quad W \propto \sqrt{N} . \tag{2.351}$$

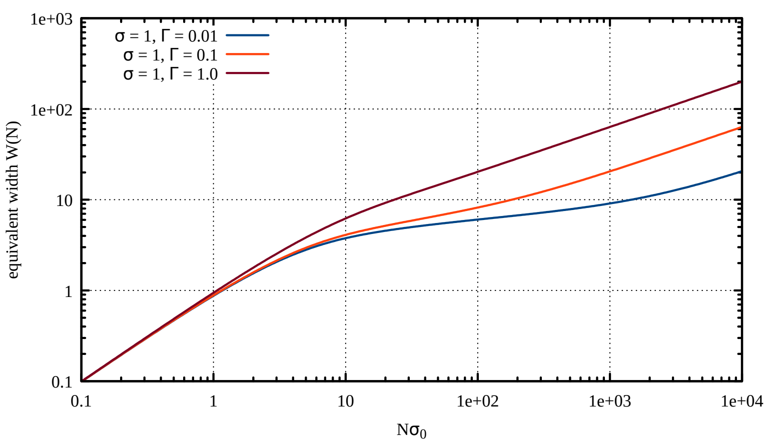

Figure 2.16 Curves-of-growth as a function of $N\sigma_0$ for three different line profiles with the same Gaussian broadening, $\sigma = 1$, but different values for the damping Γ.

Summarising, the curve-of-growth $W(N)$ behaves as

$$W(N) \propto \begin{cases} N & \text{small } N \\ \sqrt{\ln N} & \text{intermediate } N \\ \sqrt{N} & \text{large } N \end{cases} . \tag{2.352}$$

For determining the number N of absorbers, lines with different oscillator strengths f are required because then the spectral lines fall into different sections of the curve-of-growth $W(N)$ for the same N. This may prove difficult when some lines fall into the flat section of $W(N)$ where $W(N) \propto \sqrt{\ln N}$.

Problems

1. Besides their natural line width, emission lines with transition frequency ω_0 are broadened due to collisions of the emitting atoms and their thermal velocities. The collisional broadening leads to a line shape that is described by a Lorentz profile

$$\phi_{\Gamma_c}(\omega - \omega_{12}) = \frac{1}{\pi} \frac{\Gamma_c/2}{(\omega - \omega_{12})^2 + \Gamma_c^2/4} , \tag{2.353}$$

where $\Gamma_c = \sigma\langle nv \rangle$ is the collision rate, σ is the cross section for collisions, n is the number density of atoms, v their velocity and $\langle\cdot\rangle$ indicates the thermal average. The Doppler broadening leads to the Gaussian profile function

$$\phi_D = \frac{c}{\sqrt{2\pi}\omega_0\sigma_v} \exp\left[-\frac{c^2}{2\sigma_v^2}\left(\frac{\omega - \omega_0}{\omega_0}\right)^2\right] , \tag{2.354}$$

where σ_v is the velocity dispersion.

   (a) Estimate the line width for Doppler broadening from the full width at half maximum (FWHM) $\Delta\omega_D$, defined by $\phi_D(\omega_0 \pm \Delta\omega_D/2) = \phi_D(\omega_0)/2$, as a function of temperature T.
   (b) Estimate the line width $\Delta\omega_c$ due to collisions from the FWHM of $\phi_c(\omega)$ as a function of T. Assume that σ is set by the Bohr radius a_0 and that the density does not depend on temperature.
   (c) How can the results from (a) and (b) be combined to determine the density of an emitting medium?
   (d) Calculate the ratio $\Delta\omega_c/\Delta\omega_D$ for the Hα line (6563 Å) emitted from a cloud of atomic hydrogen with $n = 16\,\mathrm{cm}^{-3}$.

2.9 Radiation Quantities

In our treatment of radiation processes, we began with the classical picture of electromagnetic waves and their emission by accelerated charges. We added the photon picture when it became necessary for the treatment of momentum exchange between electromagnetic radiation and charges, and discussed quantum transitions caused by radiation. We shall proceed to discuss in this section the propagation and the transport of radiation, treating radiation in close analogy to a fluid. The main results are the definition of the specific intensity I_ω in (2.360), the angular moments (2.370) of the intensity and the demonstration that the quantity I_ω/ω^3 is relativistically invariant.

2.9.1 Specific Intensity

Let us therefore consider radiation again as a stream of particles which carry energy and momentum. In order to characterise the flow of radiation, we imagine setting up a small screen of differential area $\mathrm{d}\vec{A}$ and arbitrary orientation. Our first question is: What amount of energy is streaming per time interval $\mathrm{d}t$

into a direction enclosing the angle θ with the normal to the screen into the solid angle element $\mathrm{d}\Omega$ and within the frequency interval $\mathrm{d}\omega$?

We begin with the occupation number of photon states. Let $n_{\alpha\vec{p}}$ be the spatial number density of photons with momentum $\vec{p}$ and the polarisation state α ($\alpha = 1, 2$). The energy-momentum four-vector of a photon with energy $\hbar\omega$ is

$$p^\mu = \hbar k^\mu = \frac{\hbar\omega}{c}\begin{pmatrix} 1 \\ \hat{e} \end{pmatrix} , \tag{2.355}$$

with the unit vector $\hat{e}$ pointing into the direction of light propagation. Since photons are massless particles, the wave four-vector and hence also the four-momentum are null vectors, $\langle k, k\rangle = 0 = \langle p, p\rangle$. Therefore,

$$E = cp \quad \text{with} \quad \vec{p} = \frac{\hbar\omega}{c}\hat{e} , \quad p = \frac{\hbar\omega}{c} . \tag{2.356}$$

?

Compare (2.355) with the dispersion relation for electromagnetic waves, and (2.356) with the relativistic energy-momentum relation.

A volume element $\mathrm{d}\Gamma = \mathrm{d}^3x\mathrm{d}^3p$ of phase space is divided in cells of size $(2\pi\hbar)^3$ to account for Heisenberg's uncertainty principle: If the position of a particle is confined to $\mathrm{d}x$ in one spatial direction, its momentum in the same direction cannot be confined to better than $\mathrm{d}x\mathrm{d}p = 2\pi\hbar$. The number of cells per phase-space volume element $\mathrm{d}\Gamma$ is thus

$$\frac{\mathrm{d}^3x\mathrm{d}^3p}{(2\pi\hbar)^3} = \mathrm{d}V\frac{p^2\mathrm{d}p\mathrm{d}\Omega}{(2\pi\hbar)^3} = \mathrm{d}V\frac{\omega^2\mathrm{d}\omega\mathrm{d}\Omega}{(2\pi c)^3} , \tag{2.357}$$

where we have expressed the momentum by the frequency ω in the last step.

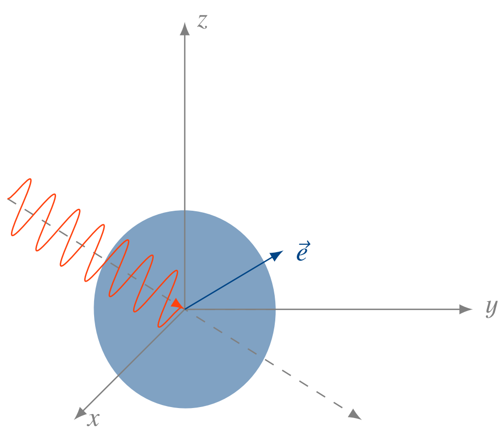

Figure 2.17 Illustration of photons streaming through an inclined area element.

The amount of energy carried by photons with momentum $\vec{p}$ through the infinitesimal screen $\mathrm{d}\vec{A}$ (Figure 2.17) is now given by the number of available phase space cells from (2.357), times the number of photons per phase-space cell with polarisation state α and momentum $\vec{p}$, times the energy $E = cp = \hbar\omega$ per photon, times the volume $\mathrm{d}V = c\mathrm{d}t\,\mathrm{d}\vec{A}\cdot\hat{e}$ covered by the screen relative to the stream of photons. Thus, we find

$$\mathrm{d}E = \frac{\omega^2\mathrm{d}\omega\mathrm{d}\Omega}{(2\pi c)^3}\sum_{\alpha=1}^{2} n_{\alpha\vec{p}}\,\hbar\omega\,\mathrm{d}A\cos\theta\,c\mathrm{d}t , \tag{2.358}$$

where θ is the angle between $\mathrm{d}\vec{A}$ and $\vec{p}$. The energy flowing through the screen per unit screen area $\mathrm{d}A$, unit time $\mathrm{d}t$ and unit frequency $\mathrm{d}\omega$ into the unit solid angle $\mathrm{d}\Omega$ defines the specific intensity I_ω of the radiation by the assignment

$$\frac{\mathrm{d}E}{\mathrm{d}t\mathrm{d}\omega\mathrm{d}A\mathrm{d}\Omega} = \sum_{\alpha=1}^{2} n_{\alpha\vec{p}} \frac{\hbar\omega^3}{(2\pi)^3 c^2} \cos\theta \equiv I_\omega \cos\theta \; . \tag{2.359}$$

For unpolarised light, $n_{1\vec{p}} = n_{2\vec{p}}$, so the sum in (2.359) merely gives a factor of two. Then, the specific intensity is related to the occupation number and the frequency by

$$I_\omega = \frac{2\hbar\omega^3}{(2\pi)^3 c^2} n_{\alpha\vec{p}} = \frac{\hbar\omega^3}{4\pi^3 c^2} n_{\alpha\vec{p}} \; . \tag{2.360}$$

Two powers of ω in the numerator are due to the volume element in phase space, the additional factor $\hbar\omega$ is the photon energy.

2.9.2 Moments of the intensity

Let us approach the intensity from a different point of view. For an electromagnetic wave in vacuum, the Poynting vector is

$$\vec{S} = \frac{c}{4\pi}\vec{E}^2\hat{k} = cU\hat{k} \; , \tag{2.361}$$

where U is the energy density. It is the energy current density in electromagnetic radiation, i.e. the electromagnetic energy flowing per unit time through unit area. Dividing by the solid angle 4π of the sphere, we find the intensity

$$I = \frac{cU}{4\pi} = \frac{1}{4\pi}\left|\vec{S}\right| \tag{2.362}$$

and its relation to the absolute magnitude of the Poynting vector. The first equation (2.362) shows that the integral of I/c over the solid angle is the energy density,

$$\int \mathrm{d}\Omega \frac{I}{c} = U \; . \tag{2.363}$$

According to the second equation (2.362), we can write the Poynting vector as $\vec{S} = 4\pi I\hat{k}$. Its integral over a sphere with arbitrary (small) radius R,

$$\int \vec{S}\cdot\mathrm{d}\vec{A} = 4\pi\int \mathrm{d}\Omega\, I\hat{k}\cdot\hat{e}_r R^2 = 4\pi R^2 \int \mathrm{d}\Omega\, I\cos\theta \tag{2.364}$$

must be the energy flowing per unit time through the sphere. Dividing by the surface area of the sphere, we find the total energy current density

$$F = \int \mathrm{d}\Omega\, I\cos\theta \tag{2.365}$$

averaged over the complete solid angle.

Maxwell's stress-energy tensor $\bar{T}$, whose components are given in (1.111), express the momentum current density. Multiplied with an oriented area element $\mathrm{d}\vec{A}$, we find the force $\mathrm{d}\vec{F} = \bar{T}\mathrm{d}\vec{A}$ exerted per area $\mathrm{d}A$ by the momentum current density since the momentum current density times an area is the momentum per unit time, hence the force. Without loss of generality, we choose $\mathrm{d}\vec{A} = \mathrm{d}A\hat{e}_z$ for an area element in the x-y plane. Since the magnetic contribution to $\bar{T}$ equals the electric contribution, we write

$$\mathrm{d}\vec{F} = \bar{T}\mathrm{d}\vec{A} = \frac{1}{2\pi}\left(\frac{\vec{E}^2}{2}\mathbb{1}_3 - \vec{E}\otimes\vec{E}\right)\mathrm{d}A\hat{e}_z = \frac{1}{2\pi}\left(\frac{\vec{E}^2}{2}\hat{e}_z - E_3\vec{E}\right)\mathrm{d}A\ . \quad (2.366)$$

It is easily seen that $\mathrm{d}F_{1,2} = 0$: There is no net force on the area element in its own plane, as expected. For $\mathrm{d}F_3$, we rather have the force per unit area, or the radiation pressure,

$$P_{\text{rad}} = \frac{\mathrm{d}F_3}{\mathrm{d}A} = \frac{1}{2\pi}\left(\frac{\vec{E}^2}{2} - E_3^2\right)\ . \quad (2.367)$$

Averaging over the solid angle, using $\langle\vec{E}^2\rangle = 3\langle E_3^2\rangle$ for a locally isotropic radiation field, writing $E_3 = E\cos\theta$, and replacing $\vec{E}^2 = 4\pi U$, we can conclude

$$P_{\text{rad}} = \frac{1}{4\pi}\int \mathrm{d}\Omega\, U\cos^2\theta = \frac{1}{4\pi}\int \mathrm{d}\Omega\, U\cos^2\theta = \frac{U}{3}\ . \quad (2.368)$$

On the one hand, this confirms the well-known result valid for all relativistic boson gases that their pressure equals a third of their energy density. On the other hand, we can substitute the intensity I from (2.362) in the first equation (2.368) to see that the radiation pressure is the second angular moment of I/c,

$$P_{\text{rad}} = \int \mathrm{d}\Omega\,\frac{I}{c}\cos^2\theta\ . \quad (2.369)$$

We have thus established the relations

$$\int \mathrm{d}\Omega\,\frac{I}{c} = U\ ,\quad F = \int \mathrm{d}\Omega\, I\cos\theta\ ,\quad P_{\text{rad}} = \int \mathrm{d}\Omega\,\frac{I}{c}\cos^2\theta \quad (2.370)$$

between the energy density U, the integrated energy current density F and the radiation pressure P_{rad} with the three lowest-order angular moments of the intensity. They will turn out to be important shortly in our discussion of radiation transport.

2.9.3 Relativistic invariance of I_ω/ω^3

Suppose now that the screen $\mathrm{d}\vec{A}$ is fixed at the origin of an unprimed coordinate frame such that it points into the $\hat{e}_z$ direction. Let it be observed from another, primed, frame moving with velocity v into the common $\hat{e}_z$ direction of the two frames. For simplicity, clocks are supposed to be synchronised such that $t = 0 = t'$ when the two frames coincide. An experimentalist resting in the unprimed frame finds by counting that

$$\mathrm{d}N = \mathrm{d}\Gamma\, n_{\vec{p}} = 2\frac{p^2\mathrm{d}p\mathrm{d}\Omega}{(2\pi\hbar)^3}\, n_{\vec{p}}\,\mathrm{d}A\, c\cos\theta\,\mathrm{d}t \quad (2.371)$$

photons have passed the screen after a time interval $\mathrm{d}t$. A fellow experimentalist resting in the primed frame counts

$$\mathrm{d}N' = 2\frac{p'^2\mathrm{d}p'\mathrm{d}\Omega'}{(2\pi\hbar)^3}\, n'_{\vec{p}'}\,\mathrm{d}A'\,(c\cos\theta' - v)\,\mathrm{d}t' \quad (2.372)$$

photons in the time interval $\mathrm{d}t'$ measured on his clock. If the two experimentalists agree to synchronise the durations of their measurements by

$$\mathrm{d}t' = \gamma \mathrm{d}t \tag{2.373}$$

to account for the relativistic time dilation, they must count the same number of photons, $\mathrm{d}N = \mathrm{d}N'$. In order to see what this implies for the occupation numbers $n_{\vec{p}}$ in the unprimed and $n'_{\vec{p}'}$ in the primed frame, we need to Lorentz transform the absolute value p of the photon momentum, the solid-angle element $\mathrm{d}\Omega$ and the direction cosine $\cos\theta$. The area elements are unchanged, $\mathrm{d}A = \mathrm{d}A'$, for they are perpendicular to the direction of the relative motion of the two frames.

We have seen in (1.45) and (1.47) that angles and solid angles change like

$$\cos\theta' = \frac{\beta + \cos\theta}{1 + \beta\cos\theta} \quad \text{and} \quad \mathrm{d}\Omega' = \frac{\mathrm{d}\Omega}{\gamma^2 (1 + \beta\cos\theta)^2} \tag{2.374}$$

under Lorentz transforms. The absolute value of the momentum is $p = E/c$, as shown in (2.356), and thus transforms like the zero component of a four-vector,

$$p' = p'^0 = \gamma\left(p^0 + \beta p^3\right) = \gamma\,(p + \beta\cos\theta p) = p\gamma(1 + \beta\cos\theta)\ . \tag{2.375}$$

We now insert the primed quantities into (2.372) to find

$$\begin{aligned}
\mathrm{d}N' &= 2\frac{p^2\mathrm{d}p\mathrm{d}\Omega}{(2\pi\hbar)^3}\frac{[\gamma(1+\beta\cos\theta)]^3}{\gamma^2(1+\beta\cos\theta)^2}\,n'_{p'}\,\mathrm{d}A\,c\left(\frac{\beta+\cos\theta}{1+\beta\cos\theta} - \beta\right)\gamma\mathrm{d}t \\
&= 2\frac{p^2\mathrm{d}p\mathrm{d}\Omega}{(2\pi\hbar)^3}\,\gamma^2(1+\beta\cos\theta)\,n'_{p'}\,\mathrm{d}A\,c\cos\theta\frac{1-\beta^2}{1+\beta\cos\theta}\,\mathrm{d}t \\
&= 2\frac{p^2\mathrm{d}p\mathrm{d}\Omega}{(2\pi\hbar)^3}\,n'_{p'}\,\mathrm{d}A\,c\cos\theta\,\mathrm{d}t
\end{aligned} \tag{2.376}$$

for the number of photons counted by the experimentalist resting in the primed frame. This agrees with $\mathrm{d}N$ from (2.371) if, and only if, the occupation numbers transform as $n_{\vec{p}} = n'_{\vec{p}'}$. With (2.360), this implies the important result that the specific intensity divided by ω^3 is invariant,

$$\frac{I_\omega}{\omega^3} = \frac{I'_\omega}{\omega'^3}\ . \tag{2.377}$$

2.10 The Planck spectrum and Einstein coefficients

In this section, the Planck spectrum is derived from first principles, i.e. from the grand-canonical partition sum of a photon gas in thermal equilibrium with a heat bath of given temperature. The first main result is the specific intensity (2.396) as a function of frequency at given temperature of thermal (black-body) radiation. Then, the Einstein coefficients for absorption, stimulated and spontaneous transition are introduced. The relations (2.425) between them required by the Planck spectrum are derived, showing that spontaneous transitions are necessary and that the rates of stimulated emission and absorption must be equal.

Example: The dipole of the Cosmic Microwave Background

To give an example, let us study an instructive consequence of the relativistic invariance of I_ω/ω^3. In its rest frame, the Cosmic Microwave Background (CMB) is an isotropic radiation field with a Planck spectrum. The occupation number is thus given by

$$n_p = \left[\exp\left(\frac{\hbar\omega}{kT_{\rm CMB}}\right) - 1\right]^{-1} , \tag{2.378}$$

where $T_{\rm CMB}$ is the CMB temperature. The energy of a photon measured by an observer moving with a four-velocity u with respect to the rest frame of the radiation is the (negative) projection of the photon's four-momentum on the four-velocity,

$$E = -\langle p, u\rangle \ . \tag{2.379}$$

This is quickly verified for an observer at rest in the rest frame of the radiation, who has $u^\mu = (c, 0)^T$ there. With p^μ from (2.355), the projection (2.379) is

$$-\langle p, u\rangle = -p^\mu u_\mu = \hbar\omega \tag{2.380}$$

for this observer, as it should be. An observer moving instead with $u' = \gamma(c, \vec{v})^T = \gamma c(1, \vec{\beta})^T$ relative to the rest frame of the radiation, however, measures the photon energy

$$E' = -\langle p, u'\rangle = \hbar\omega\gamma\left(1 - \vec{\beta}\cdot\hat{e}\right) = E\gamma\left(1 - \vec{\beta}\cdot\hat{e}\right) \ . \tag{2.381}$$

This is the relativistic Doppler shift: The moving observer measures a relative energy change of

$$\frac{E' - E}{E} = \gamma(1 - \beta\cos\theta) \tag{2.382}$$

compared to the observer at rest. For $\theta = 0$, this result simplifies to

$$\frac{E' - E}{E} = \frac{1-\beta}{\sqrt{1-\beta^2}} = \sqrt{\frac{1-\beta}{1+\beta}} \approx 1 - \beta\ , \tag{2.383}$$

where the approximation in the final step is valid for $\beta \ll 1$.
Returning to the CMB, the moving observer sees the occupation number

$$n'_{p'} = \left[\exp\left(\frac{\hbar\omega'}{kT'_{\rm CMB}}\right) - 1\right]^{-1} = \left[\exp\left(\frac{\hbar\omega\gamma(1-\beta\cos\theta)}{kT'_{\rm CMB}}\right) - 1\right]^{-1} , \tag{2.384}$$

which must be the same as n_p from (2.378). This can only be achieved if the moving observer sees a direction-dependent temperature

$$T'_{\rm CMB} = T_{\rm CMB}\gamma(1-\beta\cos\theta) \approx T_{\rm CMB}\left(1 - \frac{v}{c}\cos\theta\right) , \tag{2.385}$$

where the approximation is again valid for non-relativistic motion, $v \ll c$. The motion of the Earth relative to the rest frame of the CMB thus imprints a dipolar pattern on the measured CMB temperature (Figure 2.18). With $\beta \approx 10^{-3}$ and $T_{\rm CMB} \approx 3\,\mathrm{K}$, the amplitude of this temperature dipole is of milli-Kelvin order. ◂

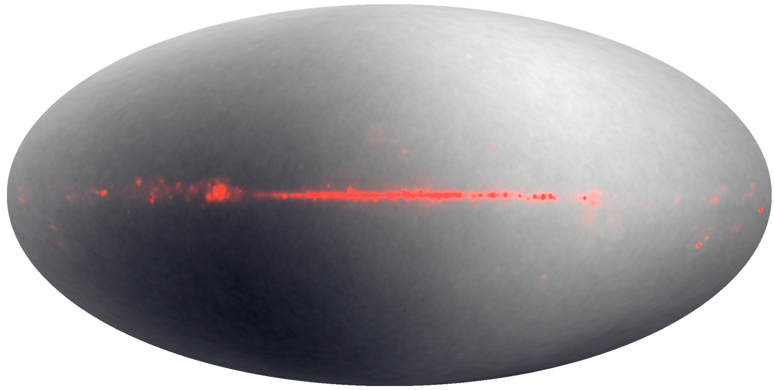

Figure 2.18 The gray-scale image shows the dipole of the Cosmic Microwave Background, measured by the Wilkinson Microwave Anisotropy Probe. In red, the emission from the Galactic disk is shown. (Provided by the WMAP Science Team)

2.10.1 The Planck spectrum

We begin by recalling some general results from statistical physics. Suppose we have an ensemble of quantum states whose occupation is in equilibrium with a heat bath of temperature T. For convenience, we shall express the temperature by the inverse thermal energy β below, $\beta = (k_\mathrm{B}T)^{-1}$. Let these states be labelled by an abstract index α which may be composed of various quantum numbers, as appropriate for the system at hand. The energies of these quantum states are called ε_α, and the quantum states are occupied n_α times. If the total number

$$N = \sum_\alpha n_\alpha \tag{2.386}$$

of occupied states is unspecified, the ensemble has the grand-canonical partition sum

$$Z_\mathrm{GC} = \sum_{N=0}^{\infty} \mathrm{e}^{\beta\mu N} \sum_{\{n_\alpha\}} \exp\left(-\beta \sum_\alpha \varepsilon_\alpha n_\alpha\right) , \tag{2.387}$$

where the summation over $\{n_\alpha\}$ is meant to indicate that the set $\{n_\alpha\}$ must obey condition (2.386). The chemical potential μ is the energy required to change the occupation number by unity. With (2.386), the partition sum (2.387) can be written

$$Z_\mathrm{GC} = \sum_{N=0}^{\infty} \sum_{\{n_\alpha\}} \exp\left[-\beta \sum_\alpha (\varepsilon_\alpha - \mu) n_\alpha\right] = \sum_{n_\alpha} \exp\left[-\beta \sum_\alpha (\varepsilon_\alpha - \mu) n_\alpha\right] , \tag{2.388}$$

where the decisive last step was possible because the sum over n_α, constrained by the fixed total occupation number N and followed by a sum over all possible values of N, amounts to an unconstrained sum over n_α. The sum in the exponential translates to a product, and we find

$$Z_\mathrm{GC} = \prod_\alpha Z_\alpha , \quad Z_\alpha = \sum_{n_\alpha} \mathrm{e}^{-\beta(\varepsilon_\alpha - \mu) n_\alpha} . \tag{2.389}$$

For Fermi-Dirac systems, $n_\alpha \in \{0, 1\}$, while $n_\alpha \in [0, \infty)$ for Bose-Einstein systems. Thus,

$$Z_\alpha^{\rm FD} = 1 + \mathrm{e}^{-\beta(\varepsilon_\alpha - \mu)} \;, \quad Z_\alpha^{\rm BE} = \frac{1}{1 - \mathrm{e}^{-\beta(\varepsilon_\alpha - \mu)}} \;, \tag{2.390}$$

where we have carried out a geometrical series for the Bose-Einstein case. The mean occupation numbers are

$$\bar{n}_\alpha = \frac{1}{Z_\alpha} \sum_\alpha n_\alpha \mathrm{e}^{-\beta(\varepsilon_\alpha - \mu) n_\alpha} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_\alpha \;. \tag{2.391}$$

Applying this to (2.390), we find

$$\bar{n}_\alpha^{\rm FD} = \frac{1}{\mathrm{e}^{\beta(\varepsilon_\alpha - \mu)} + 1} \;, \quad \bar{n}_\alpha^{\rm BE} = \frac{1}{\mathrm{e}^{\beta(\varepsilon_\alpha - \mu)} - 1} \;. \tag{2.392}$$

For a free photon gas, $\mu = 0$ because photons can spontaneously be created or destroyed. If we label photon energies ε_α by their momentum, $\varepsilon_\alpha = cp$. The energy density contained per unit photon momentum in the photon gas is then

$$\mathrm{d}U_p = 2 \cdot \frac{4\pi p^2 \mathrm{d}p}{(2\pi\hbar)^3} \cdot cp \cdot \bar{n}_p^{\rm BE} = \frac{c}{\pi^2\hbar^3} \frac{p^3 \mathrm{d}p}{\mathrm{e}^{\beta cp} - 1} \;, \tag{2.393}$$

where the factor of two accounts for the two polarisation states of each photon. Substituting the momentum p by the frequency ω through

$$E = \hbar\omega = cp \;, \tag{2.394}$$

we find the spectral energy density

$$\frac{\mathrm{d}U_\omega}{\mathrm{d}\omega} = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\mathrm{e}^{\beta\hbar\omega} - 1} \;. \tag{2.395}$$

Multiplying with $c/(4\pi)$ according to the definition of the specific intensity in (2.360), we find the Planck spectrum

$$I_\omega = \frac{c}{4\pi} \frac{\mathrm{d}U_\omega}{\mathrm{d}\omega} =: B_\omega(T) = \frac{\hbar}{4\pi^3 c^2} \frac{\omega^3}{\mathrm{e}^{\beta\hbar\omega} - 1} \;. \tag{2.396}$$

This is often expressed in terms of the frequency $\nu = \omega/(2\pi)$, for which we obtain

$$B_\nu(T) = \frac{c}{4\pi} \frac{\mathrm{d}U_\nu}{\mathrm{d}\nu} = \frac{2h}{c^2} \frac{\nu^3}{\mathrm{e}^{\beta h\nu} - 1} \;. \tag{2.397}$$

The Planck spectrum has the characteristic frequency

$$\omega_0 = \frac{k_{\rm B}T}{\hbar} \;, \quad \nu_0 = \frac{k_{\rm B}T}{h} = \frac{\omega_0}{2\pi} \;, \tag{2.398}$$

which can conveniently be used to introduce the dimension-less frequency

$$x := \frac{\omega}{\omega_0} = \frac{\nu}{\nu_0} \;, \tag{2.399}$$

in terms of which the Planck spectrum becomes (Figure 2.19)

$$B_\omega = \frac{(k_{\rm B}T)^3}{4\pi^3(\hbar c)^2} \frac{x^3}{\mathrm{e}^x - 1} \;, \quad B_\nu = \frac{2(k_{\rm B}T)^3}{(hc)^2} \frac{x^3}{\mathrm{e}^x - 1} \;. \tag{2.400}$$

The prefactors of B_ω and B_ν in (2.400) evaluate to

$$B_{\omega,0} := \frac{(k_\mathrm{B}T)^3}{4\pi^3(\hbar c)^2} = 2.12 \cdot 10^{-17}\,\frac{\mathrm{erg}}{\mathrm{cm^2\,s\,Hz\,sr}}\left(\frac{T}{\mathrm{K}}\right)^3\,, \quad B_{\nu,0} = 2\pi B_{\omega,0}\,. \tag{2.401}$$

The corresponding spectral energy density is

$$\frac{4\pi}{c}B_{\omega,0} = 8.88 \cdot 10^{-27}\,\frac{\mathrm{erg}}{\mathrm{cm^3\,Hz}}\left(\frac{T}{\mathrm{K}}\right)^3\,. \tag{2.402}$$

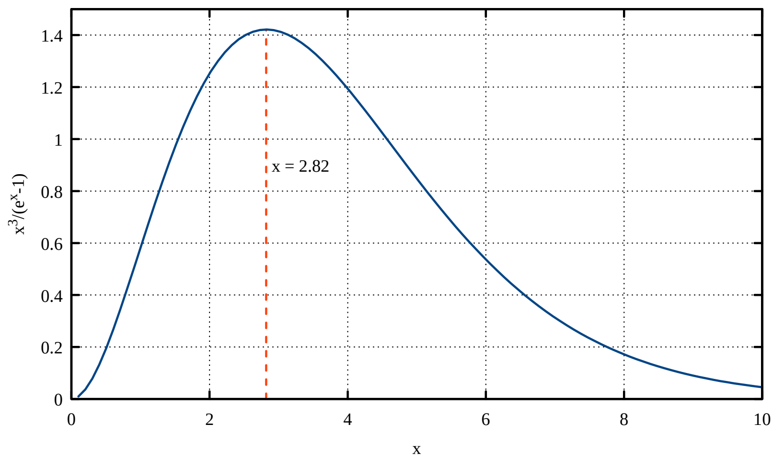

Figure 2.19 This figure shows the function $x^3/(\mathrm{e}^x - 1)$ describing the frequency dependence of the Planck spectrum. The vertical line marks the location of the maximum at $x_\mathrm{max} \approx 2.82$.

?

How would you solve an equation like (2.403)? Remind yourself of the Newton-Raphson method. Test how quickly the method converges starting with $x_0 = 3$ or $x_0 = 5$. What happens if you start with $x_0 \leq 2$?

The maximum of the Planck spectrum is located where

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{\mathrm{e}^x - 1} = 0 \quad \Rightarrow \quad (3 - x)\mathrm{e}^x = 3\,, \tag{2.403}$$

which is a transcendental equation solved by $x_\mathrm{max} \approx 2.82$. With the help of (2.403), we have

$$\frac{x_\mathrm{max}^3}{\mathrm{e}^{x_\mathrm{max}} - 1} = x_\mathrm{max}^2(3 - x_\mathrm{max}) \approx 1.43 \tag{2.404}$$

there, hence the maximum amplitude of the Planck spectrum is approximately $1.43\,B_{\omega,0}$.

For high frequencies, $x \gg 1$, the exponential in the denominator of (2.400) dominates, and the Planck spectrum can be approximated by Wien's law,

$$B_\omega \approx B_{\omega,0}\,x^3\mathrm{e}^{-x}\,, \tag{2.405}$$

while it turns into the Rayleigh-Jeans law for low frequencies, $x \ll 1$. Then, $\mathrm{e}^x - 1 \approx x$, which allows the approximation

$$B_\omega \approx B_{\omega,0}\,x^2 \tag{2.406}$$

of the spectrum. The Rayleigh-Jeans law is often used to define a radiation temperature T_rad by requiring

$$\frac{2\nu^2}{c^2}k_\mathrm{B}T_\mathrm{rad} \stackrel{!}{=} I_\nu\,. \tag{2.407}$$

Obviously, this agrees well with the thermodynamic temperature if $x \ll 1$ or $h\nu \ll 2.82\,k_\mathrm{B}T$ and $I_\nu = B_\nu$, but the deviation may become considerable for higher frequencies.

The total energy density contained in the photon ensemble is

$$U(T) = \int_0^\infty \frac{\mathrm{d}U_\omega}{\mathrm{d}\omega}\mathrm{d}\omega = \frac{(k_\mathrm{B}T)^4}{\pi^2(\hbar c)^3}\int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x - 1}\ . \tag{2.408}$$

The remaining integral is best carried out after expanding the integrand into a geometrical series,

$$\int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x - 1} = \int_0^\infty \frac{\mathrm{e}^{-x}x^3\mathrm{d}x}{1 - \mathrm{e}^{-x}} = \int_0^\infty x^3\mathrm{d}x\,\mathrm{e}^{-x}\sum_{j=0}^\infty \mathrm{e}^{-jx} = \sum_{j=1}^\infty \int_0^\infty x^3\mathrm{d}x\,\mathrm{e}^{-jx}\ . \tag{2.409}$$

Each individual integral in (2.409) gives

$$\int_0^\infty x^3\mathrm{d}x\,\mathrm{e}^{-jx} = \frac{\Gamma(4)}{j^4}\ . \tag{2.410}$$

Returning with this result to the energy density in (2.408), we find

$$U(T) = \frac{\Gamma(4)}{\pi^2}\frac{(k_\mathrm{B}T)^4}{(\hbar c)^3}\sum_{j=1}^\infty \frac{1}{j^4} = \frac{\Gamma(4)\zeta(4)}{\pi^2}\frac{(k_\mathrm{B}T)^4}{(\hbar c)^3} = \frac{\pi^2}{15}\frac{(k_\mathrm{B}T)^4}{(\hbar c)^3} =: aT^4\ , \tag{2.411}$$

where $\zeta(4) = \pi^4/90$ and $\Gamma(4) = 3! = 6$ were used in the step next to the last. Finally, the derived constant

$$a := \frac{\pi^2}{15}\frac{k_\mathrm{B}^4}{(\hbar c)^3} = 7.57 \cdot 10^{-15}\,\frac{\mathrm{erg}^4}{\mathrm{cm}^3\,\mathrm{K}^4} \tag{2.412}$$

was introduced, which is sometimes called the Stefan-Boltzmann constant. Using the same approach, we find that the number density of the photons is given by

$$n_\gamma(T) = \int_0^\infty \mathrm{d}\omega\frac{\mathrm{d}U_\omega}{\mathrm{d}\omega}\frac{1}{\hbar\omega} = \frac{(k_\mathrm{B}T)^3}{\pi^2(\hbar c)^3}\int_0^\infty \frac{x^2\mathrm{d}x}{\mathrm{e}^x - 1} = \frac{2\zeta(3)}{\pi^2}\frac{(k_\mathrm{B}T)^3}{(\hbar c)^3}\ , \tag{2.413}$$

with $\zeta(3) \approx 1.202$.

?

Carry out the integration (2.410) yourself, and confirm the result (2.413).

2.10.2 Transition Balance and the Einstein coefficients

Suppose now that we have an ensemble of simplified atoms with just two energy levels E_1 and $E_2 > E_1$ which are supposed to be in equilibrium with an ambient radiation field characterised by a temperature T. We consider the mean transition rates in an emission- and absorption process between the photons of the radiation field and transitions between the two energy levels.

Besides absorption and spontaneous emission, we will have to take stimulated emission into account, which is a consequence of the Bose character of the photons. If a quantum state is already occupied by photons, an increase in the occupation number is more likely.

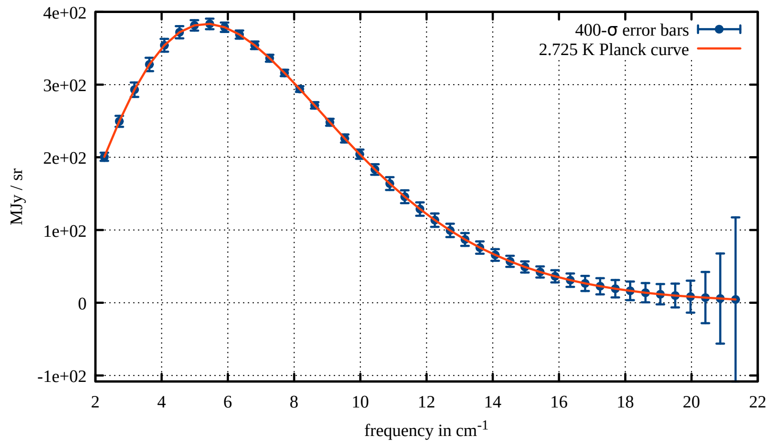

Figure 2.20 Spectrum of the Cosmic Microwave Background measured by the FIRAS instrument on-board the COBE satellite [5].

Example: The spectrum of the Cosmic Microwave Background

The best measured Planck spectrum that we know of is the spectrum of the cosmic microwave background (Figure 2.20). The CMB temperature of $T_{\mathrm{CMB}} = 2.726\,\mathrm{K}$ sets the characteristic frequency

$$\omega_{0,\mathrm{CMB}} = \frac{k_B T_{\mathrm{CMB}}}{\hbar} = 356.88 \cdot 10^9\,\mathrm{s}^{-1}\,, \quad \nu_{0,\mathrm{CMB}} = \frac{\omega_{0,\mathrm{CMB}}}{2\pi} = 56.80\,\mathrm{GHz} \tag{2.414}$$

and the frequency of the maximum specific intensity is

$$\nu_{\max,\mathrm{CMB}} = \frac{\omega_{\max,\mathrm{CMB}}}{2\pi} = 160.18\,\mathrm{GHz}\,. \tag{2.415}$$

There, the specific intensity and the spectral energy density are

$$\begin{aligned} B_{\nu_{\max,\mathrm{CMB}}} &= 1.90 \cdot 10^{-16}\,\frac{\mathrm{erg}}{\mathrm{cm}^2\,\mathrm{s\,Hz\,sr}}\,, \\ U_{\nu_{\max,\mathrm{CMB}}} &= 7.98 \cdot 10^{-26}\,\frac{\mathrm{erg}}{\mathrm{cm}^3\,\mathrm{Hz}}\,. \end{aligned} \tag{2.416}$$

The total energy density in the CMB is

$$U = 4.17 \cdot 10^{-13}\,\frac{\mathrm{erg}}{\mathrm{cm}^3}\,, \tag{2.417}$$

which is contributed by

$$n_\gamma \approx 410\,\mathrm{cm}^{-3} \tag{2.418}$$

photons per cubic centimetre. ◂

The rates of absorption and of stimulated emission, B_{12} and B_{21}, respectively, will be proportional to the specific intensity I_ω,

$$(\text{absorption rate}) \propto I_\omega B_{12} \quad \text{and} \quad (\text{stimulated emission rate}) \propto I_\omega B_{21}\,, \tag{2.419}$$

while the rate of spontaneous emission, A_{21}, will not depend on I_ω,

$$\text{(spontaneous emission rate)} \propto A_{21} \; . \tag{2.420}$$

The rate coefficients A and B are called *Einstein coefficients*.

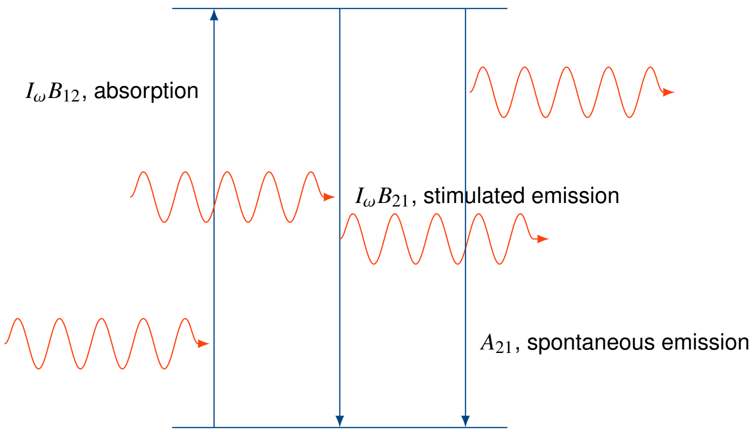

Figure 2.21 Illustration of radiative transitions between two quantum levels and the Einstein coefficients.

Now, let N_1 and N_2 be the mean occupation numbers of states with the energies E_1 and E_2. Equilibrium between transitions (Figure 2.21) will require as many transitions per unit time from E_1 to E_2 as from E_2 to E_1,

$$N_1 I_\omega B_{12} = N_2 \left[A_{21} + I_\omega B_{21}\right] \; . \tag{2.421}$$

Solving for I_ω, we see that this can be satisfied if the specific intensity is

$$I_\omega = \frac{N_2 A_{21}}{N_1 B_{12} - N_2 B_{21}} \; . \tag{2.422}$$

Since we assume thermal equilibrium, the occupation numbers N_2 and N_1 must also be related by a Boltzmann factor,

$$\frac{N_2}{N_1} = \mathrm{e}^{-\beta(E_2 - E_1)} = \mathrm{e}^{-\beta\hbar\omega} \; . \tag{2.423}$$

Inserting this into (2.422), we find

$$I_\omega = \frac{A_{21}\mathrm{e}^{-\beta\hbar\omega}}{B_{12} - B_{21}\mathrm{e}^{-\beta\hbar\omega}} = \frac{A_{21}}{B_{12}\mathrm{e}^{\beta\hbar\omega} - B_{21}} \; . \tag{2.424}$$

We can bring this into agreement with Planck's spectrum derived from quantum statistics (2.396) if, and only if, the rate coefficients satisfy Einstein's relations,

$$B_{12} = B_{21} \quad \text{and} \quad A_{21} = \frac{\hbar\omega^3}{4\pi^3 c^2} B_{21} \; . \tag{2.425}$$

This is a very interesting result, obtained by Einstein long before quantum statistics was established. It shows that without stimulated emission $B_{21} = 0$, the Planck spectrum cannot be obtained, and the microscopic rates of absorption and stimulated emission, B_{12} and B_{21} must be equal.

Problems

1. Consider an ensemble of Hydrogen atoms of temperature T. Besides the ground state, a fraction of atoms is thermally excited so that their electrons occupy higher energy levels.

   (a) Calculate the fraction n_j/n_1 of Hydrogen atoms in the excited state j relative to the ground state for $j = 2, 3$.

   (b) Calculate the relative intensity of the Lyman-β ($3 \to 1$) and Lyman-α ($2 \to 1$) lines for a cloud of atomic hydrogen with $T = 100\,\mathrm{K}$. The oscillator strengths and the wavelengths are $f_\beta = 0.0791$, $f_\alpha = 0.4162$ and $\lambda_\alpha = 1216\,\text{Å}$, $\lambda_\beta = 1026\,\text{Å}$, respectively.

2. For an ensemble of atoms with temperature T in thermal equilibrium with a radiation field, the rates for spontaneous emission A_{21}, induced emission B_{21} and absorption B_{12} between the energy levels 1 and 2 satisfy

$$N_1 I_\omega B_{12} = N_2(A_{21} + I_\omega B_{21}) , \tag{2.426}$$

   where I_ω is the specific intensity of the radiation field and $N_{1,2}$ are the numbers of atoms in the first and the second energy levels, respectively. We can use the former equation to deduce the Lyman-α cross section σ_α.

   (a) Show that the rate equation can be written as

$$3A_{21} n(\omega_\alpha) = \int \mathrm{d}\omega\, \sigma_\alpha(\omega) \frac{\omega^2}{\pi^2 c^2} n(\omega) , \tag{2.427}$$

   where $n(\omega) = [\exp(\omega/k_\mathrm{B}T) - 1]^{-1}$ is the occupation number and ω_α the circular frequency of the Lyman-α transition. The transition rate B_{12} is written in terms of the cross section σ_α.

   (b) The cross section can be written as $\sigma_\alpha = C\phi(\omega - \omega_\alpha)$ with the line profile function ϕ and a constant C. Determine C. Use $A_{21} = 6.25 \cdot 10^8\,\mathrm{s}^{-1}$ and λ_α given before. You may assume that the profile function is very narrow, i.e. it can be approximated by a Dirac delta function.

2.11 Absorption and Emission

This section begins with the definition of macroscopic coefficients for the spontaneous emission and absorption of radiation, leading to the net absorption coefficient, the opacity and the emissivity. The derivation of Kirchhoff's law (2.436) follows, which relates these quantities to the specific intensity. We then set up the radiation-transport equation and solve it under simplifying assumptions, leading to the solution (2.445). The section concludes with a discussion of continuous rather than discrete transitions.

2.11.1 Absorption coefficients and emissivity

We now want to describe how the energy transported by light is changed as the light propagates through an absorbing and emitting medium. The *absorption coefficient* α_ω is defined in terms of the energy absorbed per unit volume, time and frequency from the solid angle $\mathrm{d}^2\Omega$,

$$\alpha_\omega I_\omega = \left(\frac{\mathrm{d}E}{\mathrm{d}V\mathrm{d}t\mathrm{d}\omega\mathrm{d}^2\Omega}\right)_{\mathrm{abs}} . \tag{2.428}$$

Since the stimulated emission is also proportional to I_ω, an analogous definition applies for what is called the *induced emission coefficient*,

$$\alpha_\omega^{\mathrm{ind}} I_\omega = \left(\frac{\mathrm{d}E}{\mathrm{d}V\mathrm{d}t\mathrm{d}\omega\mathrm{d}^2\Omega}\right)_{\mathrm{ind}} . \tag{2.429}$$

To further account for the spontaneous emission, we define the *emissivity*

$$j_\omega = \left(\frac{\mathrm{d}E}{\mathrm{d}V\mathrm{d}t\mathrm{d}\omega\mathrm{d}^2\Omega}\right)_{\mathrm{spn}} , \tag{2.430}$$

which is the energy emitted spontaneously per unit volume, time and frequency into the solid-angle element $\mathrm{d}^2\Omega$. Effectively, the net absorption is the difference between absorption and stimulated emission,

$$\alpha_\omega^{\mathrm{net}} = \alpha_\omega - \alpha_\omega^{\mathrm{ind}} . \tag{2.431}$$

Since the dimension of the specific intensity I_ω is

$$\frac{\text{energy}}{\text{time} \cdot \text{area} \cdot \text{frequency} \cdot \text{solid angle}} , \tag{2.432}$$

α_ω must obviously have the dimension $(\text{length})^{-1}$. The inverse absorption coefficient α_ω^{-1} thus characterises a length, which can be identified with the mean free path for a photon of frequency ω.

Let now σ_ω be the cross section of an atom, molecule or other particle for the absorption of light of frequency ω. The number density of such absorbing particles be n, and their mass density be ρ. Then, the absorption must be due to the combined cross sections of these particles,

$$\alpha_\omega = n\sigma_\omega =: \rho\kappa . \tag{2.433}$$

The quantity κ introduced in the last step, characterising the absorption by unit mass of the medium, is called *opacity*. Its physical dimension must be an absorption cross section per unit mass, thus an area per unit mass,

$$[\kappa] = \frac{\mathrm{cm}^2}{\mathrm{g}} . \tag{2.434}$$

If the absorbing and emitting material is in equilibrium with the radiation field passing through it, the emitted and absorbed amounts of energy must equal, hence

$$j_\omega + \alpha_\omega^{\mathrm{ind}} I_\omega = \alpha_\omega I_\omega \tag{2.435}$$

or, by the definition (2.431) of the net absorption coefficient, the ratio between emissivity and net absorption coefficient must equal the specific intensity I_ω,

$$I_\omega = \frac{j_\omega}{\alpha_\omega^{\rm net}} \;, \tag{2.436}$$

which is *Kirchhoff's law*. At the same time, we have the relation (2.421) between the specific intensity and the Einstein coefficients, which must themselves be related by Einstein's relations (2.425). Combining these results, we find the relation

$$\frac{j_\omega}{\alpha_\omega^{\rm net}} = \frac{\hbar\omega^3}{4\pi^3c^2}\left(\frac{N_1}{N_2} - 1\right)^{-1} \tag{2.437}$$

between the occupation numbers N_1 and N_2 of two energy levels contributing to the radiation balance on one side, and the emissivity and the net absorption coefficient on the other. Thus, if the occupation numbers are known, the emission and absorption properties in equilibrium can be calculated, and vice versa. In particular, in thermal equilibrium between radiation and matter, the specific intensity must be given by the Planck spectrum, $I_\omega = B_\omega$, hence

$$\alpha_\omega^{\rm net} = \frac{j_\omega}{B_\omega} \;. \tag{2.438}$$

2.11.2 Radiation Transport in a Simple Case

Let us now consider an emitting and absorbing medium in which scattering can be ignored. The medium be characterised by its emissivity j_ω and a net absorption coefficient $\alpha_\omega^{\rm net}$. A light bundle passing through it has its intensity changed per unit path length by an amount

$$\mathrm{d}I_\omega = \underbrace{j_\omega\,\mathrm{d}l}_{\text{emission}} - \underbrace{\alpha_\omega^{\rm net} I_\omega\,\mathrm{d}l}_{\text{absorption}} \;, \tag{2.439}$$

from which we obtain the equation of radiation transport in its simplest case,

$$\frac{\mathrm{d}I_\omega}{\mathrm{d}l} = j_\omega - \alpha_\omega^{\rm net} I_\omega \;. \tag{2.440}$$

The homogeneous equation (2.440) is readily solved. Setting $j_\omega = 0$ for the moment,

$$\frac{\mathrm{d}I_\omega}{\mathrm{d}l} = -\alpha_\omega^{\rm net} I_\omega \quad \Rightarrow \quad \mathrm{d}\ln I_\omega = -\alpha_\omega^{\rm net}\mathrm{d}l \;, \tag{2.441}$$

thus

$$I_\omega(l) = I_{\omega,0}\,\exp\left(-\int_0^l \alpha_\omega^{\rm net}(l')\mathrm{d}l'\right) \;, \tag{2.442}$$

with an integration constant $I_{\omega,0}$ set by the incoming specific intensity.

?

Why does scattering have to be ignored for (2.440) to hold?

The inhomogeneous equation (2.440) can now be solved by a standard technique called the variation of constants. We extend the definition of the former integration constant $I_{\omega,0}$ to allow its dependence on the light path, $I_{\omega,0} = I_{\omega,0}(l)$, and find

$$j_\omega - \alpha_\omega^{\rm net} I_\omega \stackrel{!}{=} \frac{\mathrm{d}I_\omega}{\mathrm{d}l} = \left[I'_{\omega,0}(l) - I_{\omega,0}(l)\alpha_\omega^{\rm net}\right]\exp\left(-\int_0^l \alpha_\omega^{\rm net}(l')\mathrm{d}l'\right) \;, \tag{2.443}$$

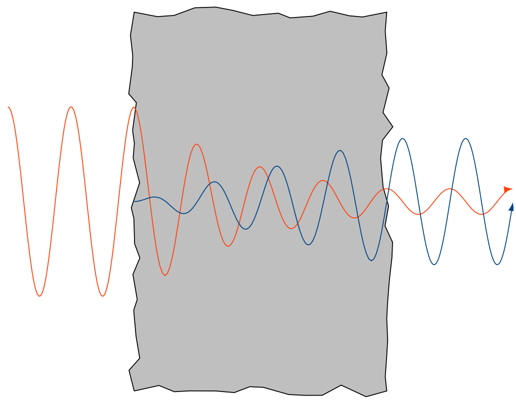

Figure 2.22 Illustration of radiation transport. While the incoming radiation is damped by the absorption, the spontaneous and the stimulated emission of the medium increase the intensity.

which, with (2.442), implies

$$j_\omega(l) = I'_{\omega,0}(l) \exp\left(-\int_0^l \alpha_\omega^{\text{net}}(l')\mathrm{d}l'\right) . \tag{2.444}$$

By separation of variables, this differential equation for $I_{\omega,0}(l)$ has the solution

$$I_{\omega,0}(l) = \int \mathrm{d}l \left[j_\omega(l) \exp\left(\int_0^l \alpha_\omega^{\text{net}}(l')\mathrm{d}l'\right)\right] + C \tag{2.445}$$

with another integration constant C set by boundary conditions.

Example: Constant emission and absorption

If $\alpha_\omega^{\text{net}}$ and j_ω are constant along the light path, the inner integral in (2.445) is simply

$$\int_0^l \alpha_\omega^{\text{net}}\mathrm{d}l = \alpha_\omega^{\text{net}} l \, , \tag{2.446}$$

while the outer integration gives

$$I_{\omega,0}(l) = \frac{j_\omega}{\alpha_\omega^{\text{net}}} \mathrm{e}^{\alpha_\omega^{\text{net}} l} + C \, . \tag{2.447}$$

By (2.442), the specific intensity then develops according to (Figure 2.22)

$$I_\omega(l) = \frac{j_\omega}{\alpha_\omega^{\text{net}}} - C\mathrm{e}^{-\alpha_\omega^{\text{net}} l} \tag{2.448}$$

along the path length of the light bundle. If, for example, the specific intensity satisfies the boundary condition $I_\omega = 0$ at $l = 0$, it changes as a function of path length like

$$I_\omega(l) = \frac{j_\omega}{\alpha_\omega^{\text{net}}} \left(1 - \mathrm{e}^{-\alpha_\omega^{\text{net}} l}\right) \, . \tag{2.449}$$

◀

Example: Radiation transport in the limiting cases of optically thick media

Interesting limiting cases of radiation transport are those of optically thin or thick media. Optically thin means that the mean free path of photons is large compared to the overall length L of the light path through the medium, $\alpha_\omega^{\text{net}} L \ll 1$, while optically thick means the opposite, $\alpha_\omega^{\text{net}} L \gg 1$. In the optically thin case, we can expand $1 - \text{e}^{-x} \approx x$ to first order and approximate

$$I_\omega(L) \approx \frac{j_\omega}{\alpha_\omega^{\text{net}}} \alpha_\omega^{\text{net}} L \approx j_\omega L \ . \tag{2.450}$$

The specific intensity is then simply the emissivity times the total path length. In the optically thick case, the exponential in (2.449) tends to zero, and

$$I_\omega(L) \approx \frac{j_\omega}{\alpha_\omega^{\text{net}}} \ . \tag{2.451}$$

This closes the loop: If the radiation is in thermal equilibrium with the optically thick medium through which it propagates, we can complete (2.451) with (2.438) to find

$$I_\omega \approx B_\omega \ . \tag{2.452}$$

This shows that radiation in thermal equilibrium with an optically thick medium leaves the medium with a Planck spectrum. ◂

Example: Radiation transport in the limiting case of optically thin media

As a further illustrative example, let us now consider optically thin, thermal emission of radio waves. As we have seen, an optically thin medium satisfies $\alpha_\omega^{\text{net}} L \ll 1$ and $I_\omega = j_\omega L$, while thermal equilibrium requires $I_\omega \approx B_\omega$. Combining these conditions, we find

$$B_\omega \approx I_\omega \approx j_\omega L = \alpha_\omega^{\text{net}} L B_\omega \ll B_\omega \ . \tag{2.453}$$

This evidently contradictory conclusion demonstrates that the two assumptions, thermal equilibrium *and* optically-thin radiation, are in manifest conflict with each other: Radiation cannot attain thermal equilibrium with an optically thin medium. ◂

Example: Planck spectrum shining through gas

For yet another instructive example, consider gas at temperature T_1 in thermal equilibrium with radiation having a Planck spectrum with temperature T_0 before it enters the gas. In the gas, Kirchhoff's law (2.436) demands

$$j_\omega = \alpha_\omega^{\text{net}} B_\omega(T_1) \tag{2.454}$$

because of the thermal equilibrium. For simplicity, we assume that T_1 and $\alpha_\omega^{\text{net}}$ are constant. In the present situation, (2.445) implies

$$\begin{aligned} I_{\omega,0}(l) &= B_\omega(T_0) + B_\omega(T_1)\alpha_\omega^{\text{net}} \int_0^l \mathrm{d}l' \mathrm{e}^{\alpha_\omega^{\text{net}} l'} \\ &= B_\omega(T_0) + B_\omega(T_1)\left(\mathrm{e}^{\alpha_\omega^{\text{net}} l} - 1\right) . \end{aligned} \tag{2.455}$$

Then, the specific intensity (2.442) is given by

$$I_\omega(l) = B_\omega(T_0)\mathrm{e}^{-\alpha_\omega^{\text{net}} l} + B_\omega(T_1)\left(1 - \mathrm{e}^{-\alpha_\omega^{\text{net}} l}\right) , \tag{2.456}$$

which is a weighed average between the two Planck spectra for temperatures T_0 and T_1. As the radiation propagates into the gas, its original Planck spectrum is gradually being replaced by the Planck spectrum determined by the gas temperature. ◂

2.11.3 Emission and Absorption in the Continuum Case

In the case of transitions between discrete energy levels, the emitted energy is determined by the number of transitions times the energy released per transition,

$$\underbrace{N_2 A_{21}}_{\text{(transition number)}} \cdot \underbrace{\hbar\omega_{12}}_{\text{(energy per transition)}} = \delta E . \tag{2.457}$$

The emissivity, defined as the energy emitted per unit time and unit volume into a unit solid angle, is thus related to the transition number by

$$j_\omega = \frac{N_2 A_{21} \hbar\omega_{12}}{4\pi} \to \frac{N_2 A_{21} \hbar\omega}{4\pi} \delta_{\mathrm{D}}(\omega - \omega_{12}) , \tag{2.458}$$

if N_2 is taken to be the occupation number of quantum states per unit volume. The Dirac delta function is introduced here for modeling a needle-sharp line transition. We generalise this last expression by replacing it with a more detailed or realistic line profile function $\phi(\omega)$,

$$j_\omega = \frac{N_2 A_{21} \hbar\omega}{4\pi} \phi(\omega) , \tag{2.459}$$

which quantifies the transition probability as a function of frequency. By a completely analogous procedure for the absorption coefficient, we find

$$\alpha_\omega = \frac{N_1 B_{12}}{4\pi} \hbar\omega \, \phi(\omega) . \tag{2.460}$$

?

Beginning with the definition of the Einstein coefficients, deduce (2.460) yourself.

Now, we consider an electron of energy E which emits the energy

$$\frac{\mathrm{d}E}{\mathrm{d}\omega \mathrm{d}t} \equiv P(\omega, E) \tag{2.461}$$

per unit time and unit frequency. Let further $f(\vec{p})$ be the momentum distribution of the electrons, then the number of electrons with energies between E and $E + \mathrm{d}E$ is

$$n(E)\,\mathrm{d}E = f(\vec{p})\,\frac{\mathrm{d}^3 p}{\mathrm{d}E}\,\mathrm{d}E = 4\pi p^2 \frac{\mathrm{d}p}{\mathrm{d}E}\,f(\vec{p})\,\mathrm{d}E \tag{2.462}$$

if we assume the electron distribution to be isotropic in momentum space. Since each electron emits the energy

$$\mathrm{d}E = P(\omega, E)\,\mathrm{d}\omega\mathrm{d}t\;, \tag{2.463}$$

we obtain the emissivity

$$j_\omega = \frac{1}{4\pi}\int_0^\infty n(E)P(\omega, E)\,\mathrm{d}E = \int_0^\infty p^2 f(p)\frac{\mathrm{d}p}{\mathrm{d}E}P(\omega, E)\,\mathrm{d}E\;. \tag{2.464}$$

By the relation (2.458) between the emissivity and the Einstein coefficient A_{21}, we have for a single transition described by the continuous line profile function $\phi(\omega)$

$$P(\omega, E_2) = \hbar\omega \int_0^{E_2} A_{21}\phi(\omega)\,\mathrm{d}E_1\;, \tag{2.465}$$

since electrons with the energy E_2 can emit through transitions to all possible states with $E_1 < E_2$. Using now Einstein's relation (2.424) between A_{21} and B_{21}, we find

$$P(\omega, E_2) = \hbar\omega\,\frac{\hbar\omega^3}{4\pi^3 c^2}\int_0^{E_2} B_{21}\phi(\omega)\,\mathrm{d}E_1\;. \tag{2.466}$$

Similarly, the net absorption coefficient is

$$\alpha_\omega = \frac{\hbar\omega}{4\pi}\int \mathrm{d}E_1 \int \mathrm{d}E_2 \left[\underbrace{n(E_1)B_{12}}_{\text{(absorption)}} - \underbrace{n(E_2)B_{21}}_{\text{(stimulated emission)}}\right]\phi(\omega)\;. \tag{2.467}$$

Exchanging the order of integrations and inserting (2.466) into the second term, that term can be rewritten as

$$\frac{\hbar\omega}{4\pi}\int \mathrm{d}E_2 n(E_2)\int \mathrm{d}E_1\,B_{21}\phi(\omega) = \frac{\pi^2 c^2}{\hbar\omega^3}\int \mathrm{d}E_2\,n(E_2)P(\omega, E_2)\;. \tag{2.468}$$

By the same procedure and using $E_2 = E_1 + \hbar\omega$, the first term can be transformed into

$$\frac{\hbar\omega}{4\pi}\int \mathrm{d}E_2 n(E_1 - \hbar\omega)\int \mathrm{d}E_1\,B_{12}\phi(\omega) = \frac{\pi^2 c^2}{\hbar\omega^3}\int \mathrm{d}E_2\,n(E_2 - \hbar\omega)P(\omega, E_2)\;. \tag{2.469}$$

We thus obtain the absorption coefficient

$$\alpha_\omega = \frac{\pi^2 c^2}{\hbar\omega^3}\int \mathrm{d}E\;[n(E - \hbar\omega) - n(E)]\;P(\omega, E)\;. \tag{2.470}$$

In thermal equilibrium with a heat bath of temperature T and far from degeneracy, the electron number density must be proportional to a Boltzmann factor,

$$n(E) \propto \exp\left(-\frac{E}{k_\mathrm{B}T}\right)\;, \tag{2.471}$$

thus the difference of the electron number densities at different energies is

$$n(E-\hbar\omega)-n(E)=n(E)\left[\exp\left(\frac{\hbar\omega}{k_{\rm B}T}\right)-1\right]\,. \tag{2.472}$$

Inserting to (2.470) with this expression, we find

$$\alpha_\omega=\frac{\pi^2c^2}{\hbar\omega^3}\left(\mathrm{e}^{\hbar\omega/k_{\rm B}T}-1\right)\int\mathrm{d}E\,n(E)\,P(\nu,E)\,. \tag{2.473}$$

The remaining integral is $4\pi j_\omega$, as (2.464) shows, allowing us to write

$$\alpha_\omega=j_\omega\,\frac{4\pi^3c^2}{\hbar\omega^3}\left(\mathrm{e}^{\hbar\omega/k_{\rm B}T}-1\right)\,. \tag{2.474}$$

A glance at (2.396) finally reveals that the factor multiplying the emissivity is the inverse Planck spectrum $B_\omega(T)$. We can thus reduce (2.474) to the relation

$$\alpha_\omega=\frac{j_\omega}{B_\omega(T)} \tag{2.475}$$

between absorption and emission, just as in the discrete case.

2.11.4 Energy transport through absorbing media

It is useful to re-write the transport equation (2.440) for radiation in spherical polar coordinates. To do so, we write the total differential $\mathrm{d}I_\omega$ of the specific intensity as

$$\mathrm{d}I_\omega=\partial_rI_\omega\mathrm{d}r+\partial_\theta I_\omega\mathrm{d}\theta \tag{2.476}$$

and use the relations

$$\mathrm{d}r=\cos\theta\mathrm{d}l\,,\quad \mathrm{d}\theta=-\frac{\sin\theta}{r}\mathrm{d}l \tag{2.477}$$

between the coordinate differentials $\mathrm{d}r,\mathrm{d}\theta$ and the path length $\mathrm{d}l$. The radiation-transport equation then reads

$$\partial_rI_\omega\cos\theta-\partial_\theta I_\omega\frac{\sin\theta}{r}=-\alpha_\omega^{\rm net}I_\omega+j_\omega\,. \tag{2.478}$$

We now integrate over frequencies ω, introduce the averaged net absorption coefficient $\bar\alpha^{\rm net}$ defined by

$$\int_0^\infty\mathrm{d}\omega\,\alpha_\omega^{\rm net}I_\omega=\bar\alpha^{\rm net}I \tag{2.479}$$

and find

$$\partial_rI\cos\theta-\partial_\theta I\frac{\sin\theta}{r}=-\bar\alpha^{\rm net}I+j\,. \tag{2.480}$$

Next, we multiply this equation by $\cos\theta/c$ and integrate over the complete solid angle $\mathrm{d}\Omega=\sin\theta\mathrm{d}\theta\mathrm{d}\varphi$. Due to the isotropy of the emissivity j, the second term on the right-hand side then vanishes altogether. The second term on the left-hand side is partially integrated to shift the derivative with respect to θ away from the intensity I. This results in

$$\partial_r\int\frac{I}{c}\cos\theta^2\,\mathrm{d}\Omega+\frac{1}{r}\int\frac{I}{c}\partial_\theta\left(\sin^2\theta\cos\theta\right)\mathrm{d}\theta\mathrm{d}\varphi=-\bar\alpha^{\rm net}\int\frac{I}{c}\cos\theta\,\mathrm{d}\Omega\,. \tag{2.481}$$

We have seen earlier in (2.370) that the angular moments of the intensity are related to the energy density U, the energy current density F and the radiation pressure $P_{\rm rad}$. Furthermore, the integral in the second term on the left-hand side of (2.481) is

$$\int \frac{I}{c}\partial_\theta\left(\sin^2\theta\cos\theta\right)\mathrm{d}\theta\mathrm{d}\varphi = \int \frac{I}{c}\left(3\cos^2\theta - 1\right)\mathrm{d}\Omega = 3P_{\rm rad} - U = 0\;, \tag{2.482}$$

leaving (2.481) in the simple form $c\partial_r P = -\bar{\alpha}^{\rm net}F$ or, with the opacity κ defined in (2.433),

$$F = -\frac{c}{\rho\kappa}\partial_r P\;. \tag{2.483}$$

The energy current density F is determined by the gradient of the radiation pressure. Since the radiation pressure P is a third of the energy density U, which is in turn given by $U = aT^4$ according to (2.411), we can write the result (2.483) in the very intuitive form

$$F = -\frac{4acT^3}{3\rho\kappa}\partial_r T\;, \tag{2.484}$$

which clearly says that the radiative energy current density through an absorbing medium is driven by the temperature gradient, and inhibited by the opacity κ.

Problems

1. The change of the specific intensity I_ω in matter per unit length is given by the radiation transport equation

$$\frac{\mathrm{d}I_\omega}{\mathrm{d}l} = j_\omega - \alpha_\omega^{\rm net}I_\omega\;, \tag{2.485}$$

where j_ω is the emissivity and $\alpha_\omega^{\rm net}$ is the net absorption coefficient. Assume that radio waves travel through a medium which has a temperature profile $T(l) = T_0\exp(-l/\lambda)$, where T_0 is the temperature at the surface and λ is a typical length scale for the temperature gradient.

   (a) Let $\alpha_\omega^{\rm net}$ be constant throughout the medium, and the radiation be in local thermal equilibrium with the medium. Solve the radiation transport equation under the condition that the incoming specific intensity at $l = 0$ is $I_{\omega,\rm in}$ and $\hbar\omega \ll k_{\rm B}T$.

   (b) Assume that the incoming spectrum is given by a power law, $I_{\omega,\rm in} = I_0\,(\omega/\omega_0)^{-\nu}$, which can be seen in many astrophysical phenomena. Determine the spectrum of the radiation once it has travelled by a distance $L \sim \lambda$ with $(\alpha_\omega^{\rm net})^{-1} \ll L$. What happens to the shape of the spectrum?

Suggested further reading: [2, 6, 7, 8, 9]

Chapter 3

Hydrodynamics

3.1 The equations of ideal hydrodynamics

In this section, the equations of ideal hydrodynamics are derived under the central assumption that the mean-free path for the particles of a fluid is infinitely small compared to all other relevant length scales. Starting point of the derivation is the Boltzmann equation from kinetic theory, moments of which are formed in a relativistically invariant way to show that the ideal hydrodynamical equations can be expressed as four-divergences of the matter-current density and of the energy-momentum tensor. The corresponding equations (3.33) are the first main result. These relativistically invariant or covariant equations are then reformulated in three-dimensional form, leading to the set of three equations (3.61) for ideal hydrodynamics: One each for the conservation of mass, momentum, and energy.

3.1.1 Particle current density and energy-momentum tensor

Even though the one-particle phase-space distribution function $f(\vec{x}, \vec{p}, t)$ is defined such that its integral over momentum space,

$$\int \mathrm{d}^3 p\, f\left(t, \vec{x}, \vec{p}\,\right) = n\left(t, \vec{x}\,\right) \tag{3.1}$$

is the spatial number density of particles, it is useful for more general considerations to derive an integral measure in momentum space that allows the construction of relativistically invariant or covariant quantities. In order to do so, let us expand the six-dimensional phase space to an eight-dimensional, extended phase space by adding time and energy as dimensions. This extended phase space is then spanned by the position and momentum four-vectors, (x^μ, p^μ), instead by their three-dimensional analogs, $(\vec{x}, \vec{p}\,)$. We denote the phase-space density in this extended phase space by $\tilde{f}(x^\mu, p^\mu)$.

Since the four components of the energy-momentum four-vector p^μ are related by the relativistic energy-momentum relation (1.66), real particles must be confined to a subspace of the extended phase space identified by the condition

$$\left(p^0\right)^2 = \vec{p}^{\,2} + m^2 c^2 \ , \tag{3.2}$$

and the condition that their total energy be positive semi-definite, $p^0 \geq 0$. At a fixed time $ct = x^0$, we must thus be able to return to the phase-space distribution function $f(t, \vec{x}, \vec{p})$ by integrating

Caution The delta distribution in (3.3) ensures that the particle system is on the hypersurface in four-dimensional momentum space defined by the energy-momentum relation (3.2). In quantum field theory, this hypersurface is called the energy shell, and the delta distribution ensures that the system is "on shell". ◀

$$\int \mathrm{d}p^0 \, \tilde{f}(x^\mu, p^\mu) \, \delta_\mathrm{D}\left[\left(p^0\right)^2 - \vec{p}^2 - m^2c^2\right] \Theta\left(p^0\right) = f(t, \vec{x}, \vec{p}) \;, \tag{3.3}$$

where the Heaviside step function $\Theta(p^0)$ ensures that the energy is non-negative.

We now use property

$$\delta_\mathrm{D}\left[g(x)\right] = \sum_i \frac{1}{|g'(x_i)|} \delta_\mathrm{D}(x_i) \tag{3.4}$$

of the Dirac delta distribution, where the sum extends over all roots x_i of $g(x)$ in the relevant domain. In the case of (3.3), $g(x)$ represents the relativistic energy-momentum relation. It has two roots in total, one of them positive, hence

$$\delta_\mathrm{D}\left[\left(p^0\right)^2 - \vec{p}^2 - m^2c^2\right] = \frac{1}{2p^0} \delta_\mathrm{D}\left(p^0 - p^0_E\right) \;, \tag{3.5}$$

where p^0_E on the right-hand side is related to the particle energy by $cp^0_E = E$. Returning with this result to the integral in (3.3), we see that we can write

$$\begin{aligned} &\int \mathrm{d}^4p \, \tilde{f}(x^\mu, p^\mu) \, \delta_\mathrm{D}\left[\left(p^0\right)^2 - \vec{p}^2 - m^2c^2\right] \Theta\left(p^0\right) \\ &\qquad = \frac{c}{2} \int \frac{\mathrm{d}^3p}{E} \tilde{f}\left(x^\mu, p^0 = p^0_E, \vec{p}\right) \;. \end{aligned} \tag{3.6}$$

A further integration over d^4x must return the total number of particles,

$$\begin{aligned} N &= \int \mathrm{d}^4x \mathrm{d}^4p \, \tilde{f}(x^\mu, p^\mu) \, \delta_\mathrm{D}\left[\left(p^0\right)^2 - \vec{p}^2 - m^2c^2\right] \Theta\left(p^0\right) \\ &= c \int \mathrm{d}^4x \int \frac{\mathrm{d}^3p}{2E} \tilde{f}\left(x^\mu, p^0 = p^0_E, \vec{p}\right) \end{aligned} \tag{3.7}$$

which must be Lorentz invariant. The four-dimensional volume elements d^4x and d^4p are both relativistically invariant because Lorentz transforms have unit determinant. Since the Dirac-delta distribution and the Heaviside step function in (3.7) are manifestly Lorentz invariant, we conclude that the distribution function $\tilde{f}$ in the extended phase-space must be Lorentz invariant as well. The second equality in (3.7) then shows that d^3p/E is a Lorentz-invariant integral measure for integrations over three-dimensional momentum space. The one-particle distribution function $\tilde{f}(x^\mu, p^\mu)$ in extended phase space, constrained by the condition $p^0 = p^0_E = E/c$, can be identified with the distribution function $f(t, \vec{x}, \vec{p})$ in ordinary phase space, which is therefore also a Lorentz invariant.

Armed with this important insight, we now define two Lorentz-covariant quantities, a four-vector

$$J^\alpha(t, \vec{x}) := c \int \frac{\mathrm{d}^3p}{E} f(t, \vec{x}, \vec{p}) \; p^\alpha \tag{3.8}$$

and a rank-2 tensor

$$T^{\alpha\beta}(t, \vec{x}) := c^2 \int \frac{\mathrm{d}^3p}{E} f(t, \vec{x}, \vec{p}) \; p^\alpha p^\beta \;. \tag{3.9}$$

With $p^0 = E/c$ and the further relations $p^i = \gamma m v^i = E v^i/c^2$, we can write the components of J^α as

$$J^0(t,\vec{x}) = n(t,\vec{x}) \;, \quad J^i = \frac{1}{c}\int \mathrm{d}^3p\, f(t,\vec{x},\vec{p})\;\dot{x}^i = \frac{n(t,\vec{x})\left\langle\dot{x}^i\right\rangle}{c} \;, \tag{3.10}$$

where we have used in the final step that arbitrary properties Q of the system considered can be averaged over momenta by the operation

$$\langle Q\rangle(t,\vec{x}) = \frac{\int \mathrm{d}^3p\, Q\, f(t,\vec{x},\vec{p})}{\int \mathrm{d}^3p\, f(t,\vec{x},\vec{p})} = \frac{\int \mathrm{d}^3p\, Q\, f(t,\vec{x},\vec{p})}{n(t,\vec{x})} \;. \tag{3.11}$$

The quantity $\langle\dot{x}^i\rangle$ introduced in (3.10) above is therefore the i component of the velocity averaged over all particles near position $\vec{x}$ at time t. We denote this mean velocity by

$$\vec{v} = \vec{v}(t,\vec{x}) = \left\langle\dot{\vec{x}}\right\rangle(t,\vec{x}) \tag{3.12}$$

and write the four-vector J^α as

$$J^\alpha = \frac{n(t,\vec{x})}{c}\begin{pmatrix} c \\ \vec{v}\end{pmatrix} \;. \tag{3.13}$$

It characterises the particle current density.

Turning now to the tensor components $T^{\alpha\beta}$, we find by using $p^0 = E/c = \gamma m c$ and $p^i = \gamma m \dot{x}^i$ that

$$T^{00} = mc^2 \int \mathrm{d}^3p\, f(t,\vec{x},\vec{p})\;\gamma = mn(t,\vec{x})\,c^2\,\langle\gamma\rangle = \rho(t,\vec{x})\,c^2\langle\gamma\rangle \;, \tag{3.14}$$

where the mass density $\rho(t,\vec{x}) = mn(t,\vec{x})$ was identified, further

$$T^{0i} = \rho(t,\vec{x})\,c\left\langle\gamma\dot{x}^i\right\rangle \quad \text{and} \quad T^{ij} = \rho(t,\vec{x})\left\langle\gamma\dot{x}^i\dot{x}^j\right\rangle \;. \tag{3.15}$$

Their meaning becomes perhaps most evident in the non-relativistic limit. Then, we can Taylor-expand the Lorentz factor γ to lowest order,

$$\gamma \approx 1 + \frac{\beta^2}{2} \;, \quad \langle\gamma\rangle \approx 1 + \frac{1}{2c^2}\left\langle\dot{\vec{x}}^2\right\rangle \;, \tag{3.16}$$

and the time-time element T^{00} turns into

$$T^{00} \approx \rho c^2 + \frac{\rho}{2}\left\langle\dot{\vec{x}}^2\right\rangle \;, \tag{3.17}$$

which is the sum of the rest-mass and the kinetic energy densities of the particle ensemble near position $\vec{x}$ at time t. In this way, the tensor $T^{\alpha\beta}$ turns out to be the energy-momentum tensor of the ensemble.

To third order in v/c, we can approximate the time-space components of the energy-momentum tensor by

$$T^{0i} \approx \rho c v^i + \frac{\rho}{2c}\left\langle\dot{\vec{x}}^2\dot{x}^i\right\rangle \;, \quad T^{ij} \approx \rho\left\langle\dot{x}^i\dot{x}^j\right\rangle \;. \tag{3.18}$$

The first term in T^{0i} is the rest-energy current density, while the expression

$$\frac{\rho}{2}\left\langle\dot{\vec{x}}^2\dot{x}^i\right\rangle =: q^i \tag{3.19}$$

in the second term is the mean flow of kinetic energy.

3.1.2 Collisional invariants and the fluid approximation

We now return to the Boltzmann equation (1.155) and exclude external, macroscopic forces for now. This allows us to set $\dot{\vec{p}} = 0$ and write

$$\partial_t f(t, \vec{x}, \vec{p}) + \dot{\vec{x}} \cdot \vec{\nabla} f(t, \vec{x}, \vec{p}) = C[f] \; . \tag{3.20}$$

Our next concern is the collision term on the right-hand side, which is yet unspecified.

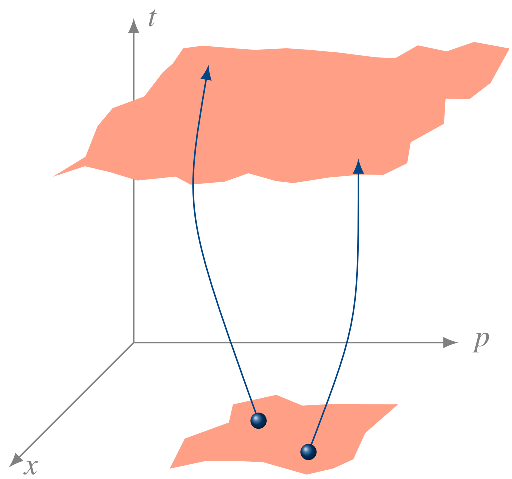

Figure 3.1 Illustration of Liouville's theorem: Trajectories of classical particles are not lost in phase space.

Recall how Boltzmann's equation was derived earlier from Liouville's equation (cf. Figure 3.1). We closed the BBGKY hierarchy by the assumption that the two-particle distribution function could be factorised into one-particle contributions. In other words, collisions between fluid particles were restricted to two-body collisions of otherwise independent particles. We can make substantial progress now by limiting our consideration to collisional invariants. These are defined to be quantities whose sum is conserved in each of these two-body collisions. If the particles can be treated as unstructured, solid bodies without internal degrees of freedom, then the particle number, their total energy and momentum can be considered conserved. Summing over many particles undergoing many collisions, none of these collisional invariants can be changed. We can thus expect that the integrals

$$\int \mathrm{d}^3 p \, C[f] \quad \text{and} \quad \int \mathrm{d}^3 p \, C[f] p^\mu \tag{3.21}$$

must vanish if their integration domains in momentum-space are chosen such that many collisions are contained. To make this possible is the essential motivation for the basic assumption underlying hydrodynamics.

A fluid in the sense of hydrodynamics is an ensemble of many particles whose mean-free path λ is very short compared to all other relevant length scales. Let the overall scale of the system be L, and the scale on which the system's

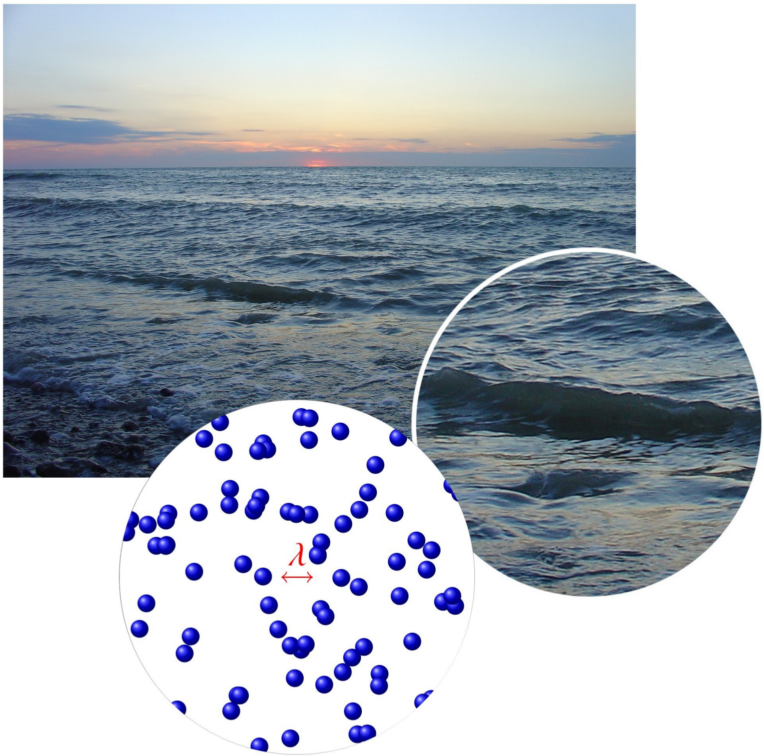

Figure 3.2 Illustration of the fundamental assumption of hydrodynamics: Collections of particles can be treated as a fluid if their mean-free path λ is very much smaller than the typical scale l on which macroscopic properties change, which is in turn much smaller than the overall scale L of the system.

macroscopic physical properties are to be determined be l. Then, for the system to be a fluid, it must be possible to establish the hierarchy of scales

$$\lambda \ll l \ll L\,. \tag{3.22}$$

A swimming pool sets a good example (see also Figure 3.2). For the overall scale, we can take the smallest of its three dimensions length, width and depth, which will be of the order of a metre. If we want to describe the flow of the water in the pool, we need to know its physical properties, such as its local flow velocity, on a length scale of perhaps a millimetre. Under normal conditions, a cubic millimetre of water will weigh 10^{-3} g. Since the mass of a single water molecule is 18 atomic mass units or $3 \cdot 10^{-23}$ g, there are $\sim 3 \cdot 10^{19}$ water molecules in each cubic millimetre, with a mean inter-particle separation of $\sim 3 \cdot 10^{-8}$ cm. The mean-free path is certainly smaller than this, so the hydrodynamical conditions are clearly satisfied very comfortably.

Given this fundamental assumption underlying hydrodynamics, we may safely assert that even a small spatial subvolume of the fluid will contain very many particles. They undergo frequent two-body collisions, in each of which five collisional invariants are conserved: the total particle number, the energy and the momentum. Any individual two-particle collision may or may not change the number of particles in a given phase-space cell. Averaging over an increasing number of collisions, however, the net change in the number of particles, their energies and momenta will decrease since all of these quantities must be conserved. The fundamental assumption of hydrodynamics assures that an average over very many collisions is possible even if the volume is small over which the average is extended.

We can thus conclude that, by the assumption (3.22) defining a fluid, the five integrals

$$\int \mathrm{d}^3p\, C[f] \quad \text{and} \quad \int \mathrm{d}^3p\, C[f]p^\mu \tag{3.23}$$

over the collision term all vanish.

We now return to the force-free Boltzmann equation (3.20) and take its lowest-order moments by carrying out the integrals given in (3.23). The lowest-order moment is

$$\partial_t n\,(t,\vec{x}) + \int \mathrm{d}^3p\, \dot{\vec{x}} \cdot \vec{\nabla} f(t,\vec{x},\vec{p}\,) = 0\ . \tag{3.24}$$

Since $\vec{v}$ and $\vec{x}$ are independent, the spatial gradient applied to $f(t,\vec{x},\vec{p}\,)$ can be pulled out of the integral, giving

$$\partial_t n\,(t,\vec{x}) + \vec{\nabla} \cdot \int \mathrm{d}^3p\, \dot{\vec{x}}\, f\,(t,\vec{x},\vec{p}\,) = 0\ . \tag{3.25}$$

Comparing this equation with (3.10), we see that we can rewrite it in terms of the four-vector J^α for the particle current density in the very simple, manifestly covariant and Lorentz-invariant form

$$\partial_\alpha J^\alpha = 0\ . \tag{3.26}$$

Next, we form the higher order moments of the force-free Boltzmann equation. This means that we multiply it with p^μ and integrate over d^3p. Beginning with p^0, we first find

$$\partial_t \int \mathrm{d}^3p\, f\,(t,\vec{x},\vec{p}\,)\ p^0 + \partial_i \int \mathrm{d}^3p\, f\,(t,\vec{x},\vec{p}\,)\ \dot{x}^i p^0 = 0\ . \tag{3.27}$$

Recalling $p^0 = E/c$ and $\dot{x}^i = p^i c^2/E$, further using $\partial_t = c\partial_0$, we can bring this equation into the form

$$c^2\partial_0 \int \frac{\mathrm{d}^3p}{E}\, f\,(t,\vec{x},\vec{p}\,)\ p^0p^0 + c^2\partial_i \int \frac{\mathrm{d}^3p}{E}\, f\,(t,\vec{x},\vec{p}\,)\ p^0p^i = 0\ . \tag{3.28}$$

Here, we can identify the time-time and time-space components of the energy-momentum tensor defined in (3.9) and bring (3.28) into the covariant form

$$\partial_\mu T^{0\mu} = 0\ . \tag{3.29}$$

Finally, we multiply the force-free Boltzmann equation with p^j to obtain

$$\partial_t \int \mathrm{d}^3p\, f\,(t,\vec{x},\vec{p}\,)\ p^j + \partial_i \int \mathrm{d}^3p\, f\,(t,\vec{x},\vec{p}\,)\ \dot{x}^i p^j = 0\ . \tag{3.30}$$

Again, we insert a factor $1 = cp^0/E$ into the first term and use $\dot{x}^i = p^i c^2/E$ in the second to write this equation as

$$c^2\partial_0 \int \frac{\mathrm{d}^3 p}{E} f\left(t, \vec{x}, \vec{p}\right)\, p^0 p^j + c^2\partial_i \int \frac{\mathrm{d}^3 p}{E} f\left(t, \vec{x}, \vec{p}\right)\, p^i p^j = 0\,, \tag{3.31}$$

which can be summarised as

$$\partial_\mu T^{j\mu} = 0\,. \tag{3.32}$$

We thus arrive at the very important and intuitive result that, under the fundamental assumption of hydrodynamics, the zeroth- and first-order moments of the force-free Boltzmann equation can be written as

$$\boxed{\partial_\mu J^\mu = 0\,, \quad \partial_\mu T^{\mu\nu} = 0}\,, \tag{3.33}$$

with the four-vector J^μ of the particle-current density and the energy-momentum tensor $T^{\mu\nu}$ of the particle ensemble. These five equations express the conservation of particles, energy and momentum and can already be seen as one form of the hydrodynamical equations.

Recall the assumptions their derivation was based upon. Besides the fundamental assumption (3.22) of hydrodynamics, we made use of five collisional invariants to argue that the momentum-space integrals over the collision term $C[f]$ should vanish. These were the total particle number, their energies and momenta. If any of these assumptions is violated, the conservation equations (3.33) cannot hold any longer. For example, the particle number may change in collisions if particles combine to form molecules. The (kinetic) energy need not be conserved if internal degrees of freedom in the particles can be excited in collisions. Under such circumstances, one needs to return to the collisional Boltzmann equation and work out the collision term explicitly.

The manifestly Lorentz-covariant equations (3.33) can easily be ported into General Relativity. We simply need to replace the partial by covariant derivatives,

$$\nabla_\mu J^\mu = 0\,, \quad \nabla_\mu T^{\mu\nu} = 0 \tag{3.34}$$

to find the fundamental equations of generally-relativistic hydrodynamics.

3.1.3 The equations of ideal hydrodynamics

We now insert the specific expressions (3.13) for the components of the particle-current density J^α as well as the non-relativistic approximations (3.17) and (3.18) for the components of the energy-momentum tensor $T^{\mu\nu}$ into the general conservation equations (3.33). For the particle-current density, we find immediately

$$\partial_t n\left(t, \vec{x}\right) + \vec{\nabla} \cdot \left[n\left(t, \vec{x}\right)\vec{v}\right] = 0\,. \tag{3.35}$$

Multiplying with the particle mass m turns the number density $n\left(t, \vec{x}\right)$ into the mass density $\rho(t, \vec{x})$, which then satisfies the equation

$$\partial_t \rho + \vec{\nabla} \cdot (\rho\vec{v}) = 0\,. \tag{3.36}$$

This is the continuity equation, or the equation for mass conservation: The local density ρ changes with time by the divergence of the matter current density $\rho\vec{v}$.

In the conservation equation $\partial_\mu T^{\mu\nu} = 0$, the time component, $\nu = 0$, selects the energy-conservation equation, while momentum conservation is expressed by its spatial components, $\nu = i$. With the non-relativistic approximations for T^{00} and T^{0i} and T^{ij} derived in (3.17) and (3.18), we find

$$c^{-1}\partial_t\left(\rho c^2 + \frac{\rho}{2}\left\langle\dot{\vec{x}}^2\right\rangle\right) + \vec{\nabla}\cdot\left(\rho c\vec{v} + \frac{\vec{q}}{c}\right) = 0 \tag{3.37}$$

for the conservation of the energy density, and

$$c^{-1}\partial_t\left(\rho c\vec{v} + \frac{\vec{q}}{c}\right) + \vec{\nabla}\cdot\left(\rho\left\langle\dot{\vec{x}}\otimes\dot{\vec{x}}\right\rangle\right) = 0 \tag{3.38}$$

for momentum conservation. Recall that the vector $\vec{q}$ is the current density of the kinetic energy, defined in (3.19). We can re-arrange the energy-conservation equation (3.37) to read

$$c\left[\partial_t\rho + \vec{\nabla}\cdot(\rho\vec{v})\right] + c^{-1}\left[\partial_t\left(\frac{\rho}{2}\left\langle\dot{\vec{x}}^2\right\rangle\right) + \vec{\nabla}\cdot\vec{q}\right] = 0\ . \tag{3.39}$$

By the continuity equation (3.36), the first term in brackets vanishes, which expresses the fact that mass conservation implies the conservation of rest-mass energy. The energy-conservation equation is thus simplified to

$$\partial_t\left(\frac{\rho}{2}\left\langle\dot{\vec{x}}^2\right\rangle\right) + \vec{\nabla}\cdot\vec{q} = 0\ . \tag{3.40}$$

Comparing terms in the momentum-conservation equation (3.38), we see that the current density of the kinetic energy $\vec{q}$ is smaller by a factor of order v^2/c^2 compared to the current density $\rho c^2\vec{v}$ of the rest-energy density. We can thus safely neglect it in our non-relativistic approximation and write momentum conservation as

$$\partial_t\left(\rho v^i\right) + \vec{\nabla}\cdot\left(\rho\left\langle\dot{\vec{x}}\otimes\dot{\vec{x}}\right\rangle\right) = 0\ . \tag{3.41}$$

Having arrived at this point, we split up the microscopic velocities $\dot{\vec{x}}$ into the mean macroscopic velocity $\vec{v}$ of the fluid flow and a random velocity $\vec{u}$ about the mean,

$$\dot{\vec{x}} = \vec{v} + \vec{u}\ . \tag{3.42}$$

As $\vec{v}$ has been defined as the average over $\dot{\vec{x}}$, the average of $\vec{u}$ must vanish by definition. The average over the squared microscopic velocity is therefore

$$\left\langle\dot{\vec{x}}^2\right\rangle = \vec{v}^2 + \left\langle\vec{u}^2\right\rangle\ , \tag{3.43}$$

which allows us to split up the kinetic energy density into a macroscopic part $\rho v^2/2$ and a microscopic or internal part $\rho\langle u^2\rangle/2$. If this internal kinetic energy density is of thermal origin, we can identify it with the thermal energy density

$$\varepsilon = \frac{\rho}{2}\left\langle u^2\right\rangle = \frac{3}{2}nk_\mathrm{B}T\ . \tag{3.44}$$

The kinetic-energy current density $\vec{q}$ has been introduced as the average

$$\vec{q} = \frac{\rho}{2}\left\langle\dot{\vec{x}}^2\dot{\vec{x}}\right\rangle \tag{3.45}$$

in (3.19). Splitting the microscopic velocities as in (3.42), we can write

$$\left\langle \dot{\vec{x}}^2 \dot{\vec{x}} \right\rangle = \left\langle \left(v^2 + 2\vec{v}\cdot\vec{u} + u^2\right)(\vec{v}+\vec{u})\right\rangle = v^2\vec{v} + \left\langle u^2 \right\rangle \vec{v} + 2\left\langle \vec{u}\otimes\vec{u}\right\rangle \vec{v} \tag{3.46}$$

because all terms must vanish in which components of $\vec{u}$ appear linearly. Thus, the kinetic-energy current density is

$$\vec{q} = \frac{\rho}{2}\left(v^2 + \left\langle u^2 \right\rangle\right)\vec{v} + \rho\left\langle \vec{u}\otimes\vec{u}\right\rangle\vec{v} = \left(\frac{\rho}{2}v^2 + \varepsilon\right)\vec{v} + \rho\left\langle \vec{u}\otimes\vec{u}\right\rangle\vec{v}\,. \tag{3.47}$$

The first two terms are the current densities of the macroscopic and the internal kinetic energies, and the meaning of the third term remains to be clarified.

We finally study the stress-energy tensor $\bar{T}$ with elements T^{ij},

$$\bar{T} = \rho\left\langle \dot{\vec{x}}\otimes\dot{\vec{x}}\right\rangle = \rho\left\langle(\vec{v}+\vec{u})\otimes(\vec{v}+\vec{u})\right\rangle = \rho\left(\vec{v}\otimes\vec{v} + \left\langle\vec{u}\otimes\vec{u}\right\rangle\right)\;, \tag{3.48}$$

where we have used once more that all terms linear in $\vec{u}$ must average to zero. The average $\langle\vec{u}\otimes\vec{u}\rangle$ appears again. In the rest frame of the macroscopic fluid flow, $\vec{v} = 0$. The trace of the stress-energy tensor is then three times the pressure of the fluid,

$$\rho\,\mathrm{Tr}\left\langle\vec{u}\otimes\vec{u}\right\rangle = \rho\left\langle u^2\right\rangle = 3P\,. \tag{3.49}$$

If the fluid is microscopically isotropic, the random velocity components u^i must be independent, hence $\langle u^i u^j\rangle = 0$ for $i \neq j$ and

$$\rho\left\langle u^i u^i\right\rangle = \frac{\rho}{3}\,\mathrm{Tr}\left\langle\vec{u}\otimes\vec{u}\right\rangle = P\,. \tag{3.50}$$

Combining these arguments, we can write

$$\vec{q} = \left(\frac{\rho}{2}v^2 + \varepsilon + P\right)\vec{v} \quad\text{and}\quad \bar{T} = \rho\vec{v}\otimes\vec{v} + P\mathbb{1}_3\,. \tag{3.51}$$

With these results, we can now bring the momentum-conservation equation (3.41) into the form

$$\partial_t(\rho\vec{v}) + \vec{\nabla}\cdot(\rho\vec{v}\otimes\vec{v}) + \vec{\nabla}P = 0\,. \tag{3.52}$$

Once more, we can re-group terms suitably to identify and remove the two terms representing mass conservation,

$$\left[\partial_t\rho + \vec{\nabla}\cdot(\rho\vec{v})\right]\vec{v} + \rho\left[\partial_t\vec{v} + \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v}\right] + \vec{\nabla}P = \rho\left(\partial_t + \vec{v}\cdot\vec{\nabla}\right)\vec{v} + \vec{\nabla}P = 0\,. \tag{3.53}$$

Momentum conservation is thus expressed by Euler's equation

$$\rho\left(\partial_t + \vec{v}\cdot\vec{\nabla}\right)\vec{v} + \vec{\nabla}P = 0\,. \tag{3.54}$$

The differential operator in parentheses is the total time derivative,

$$\partial_t + \vec{v}\cdot\vec{\nabla} = \partial_t + \frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial}{\partial\vec{x}} = \frac{\mathrm{d}}{\mathrm{d}t}\,. \tag{3.55}$$

Equation (3.54) thus simply states that ideal fluids are accelerated by pressure gradients,

$$\rho\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\vec{\nabla}P\,, \tag{3.56}$$

in absence of external, macroscopic forces.

We finally turn to the energy-conservation equation (3.40). With our results (3.43) and (3.51), it becomes

$$\partial_t\left(\frac{\rho}{2}v^2+\varepsilon\right)+\vec{\nabla}\cdot\left[\left(\frac{\rho}{2}v^2+\varepsilon+P\right)\vec{v}\right]=0\;. \tag{3.57}$$

Expanding the derivatives and re-grouping terms, we can identify those terms here that must vanish due to mass conservation and momentum conservation,

$$\frac{v^2}{2}\left[\partial_t\rho+\vec{\nabla}\cdot(\rho\vec{v})\right]+\frac{\rho}{2}\left(\partial_t+\vec{v}\cdot\vec{\nabla}\right)v^2+\partial_t\varepsilon+\vec{\nabla}\cdot\left[(\varepsilon+P)\vec{v}\right]=0\;. \tag{3.58}$$

By mass conservation, the first term in brackets vanishes. By momentum conservation, the second term in parentheses is

$$\frac{\rho}{2}\left(\partial_t+\vec{v}\cdot\vec{\nabla}\right)v^2=\rho\vec{v}\cdot\left(\partial_t+\vec{v}\cdot\vec{\nabla}\right)\vec{v}=-\vec{v}\cdot\vec{\nabla}P=-\vec{\nabla}\cdot(P\vec{v})+P\vec{\nabla}\cdot\vec{v}\;. \tag{3.59}$$

With this identification, the energy-conservation equation shrinks to

$$\partial_t\varepsilon+\vec{\nabla}\cdot(\varepsilon\vec{v})+P\vec{\nabla}\cdot\vec{v}=0\;. \tag{3.60}$$

Again, this has a very intuitive interpretation: The internal energy density changes locally not only by the current density $\varepsilon\vec{v}$, but also by the pressure-volume work $P\vec{\nabla}\cdot\vec{v}$ that the fluid has to exert against its surroundings. If the velocity field is divergent, $\vec{\nabla}\cdot\vec{v}>0$, the fluid expands, and part of its internal energy must be used for working against the pressure of its surroundings. Conversely, if $\vec{\nabla}\cdot\vec{v}<0$, the velocity field is convergent, the fluid is compressed, and its surroundings increase its internal energy by pressure-volume work.

Summarising, our final set of equations for ideal hydrodynamics reads

$$\begin{aligned}\partial_t\rho+\vec{\nabla}\cdot(\rho\vec{v})=0\;,\\ \partial_t\vec{v}+\left(\vec{v}\cdot\vec{\nabla}\right)\vec{v}+\frac{\vec{\nabla}P}{\rho}=0\;,\\ \partial_t\varepsilon+\vec{\nabla}\cdot(\varepsilon\vec{v})+P\vec{\nabla}\cdot\vec{v}=0\;.\end{aligned} \tag{3.61}$$

They express mass, momentum, and energy conservation in a very intuitive way. They are five equations for the mass density ρ, the internal energy density ε, the pressure P, and the velocity $\vec{v}$, which are six quantities in total. The set (3.61) of equations thus needs to be complemented by an equation of state that relates the pressure to the density, $P=P(\rho)$. The second equation, describing momentum conservation, is often called Euler's equation.

With a slight rearrangement in the energy-conservation equation, we can identify the total time derivative of the energy density,

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}t}+(\varepsilon+P)\vec{\nabla}\cdot\vec{v}=0\;. \tag{3.62}$$

From the point of view of thermodynamics, this is quite intuitive since the sum of the internal energy density ε and the pressure P is the enthalpy per unit volume, or the enthalpy density h,

$$h=\varepsilon+P\;. \tag{3.63}$$

Energy conservation can thus also be expressed by

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}t} + h\vec{\nabla} \cdot \vec{v} = 0 \,, \tag{3.64}$$

which is the first law of thermodynamics at given pressure.

If external, macroscopic forces are present, such as the gravitational force, the momentum-conservation equation must be augmented by the corresponding force densities. Let Φ be the Newtonian gravitational potential, its negative gradient $-\vec{\nabla}\Phi$ is the gravitational force per unit mass. It can be added to the right-hand side of the momentum-conservation equation to yield

$$\partial_t \vec{v} + \left(\vec{v} \cdot \vec{\nabla}\right) \vec{v} + \frac{\vec{\nabla} P}{\rho} = -\vec{\nabla}\Phi \,. \tag{3.65}$$

It is sometimes useful to write the complete set of equations (3.61) in terms of total time derivatives. It then reads

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho\vec{\nabla} \cdot \vec{v} = 0 \,, \quad \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} + \frac{\vec{\nabla} P}{\rho} = -\vec{\nabla}\Phi \,, \quad \frac{\mathrm{d}\varepsilon}{\mathrm{d}t} + h\vec{\nabla} \cdot \vec{v} = 0 \,. \tag{3.66}$$

Problems

1. The energy-momentum tensor is defined as

$$T^{\mu\nu} \equiv c^2 \int \frac{\mathrm{d}^3 p}{E(p)} \, p^\mu p^\nu f(\vec{x}, \vec{p}, t) \,, \tag{3.67}$$

where $(p^\mu) = (E/c, \vec{p})^T$ is the four-momentum, E the energy, and $f(\vec{x}, \vec{p}, t)$ the one-particle phase-space density distribution. While the energy density is $\varepsilon = T^{00}$, the pressure is given by one third of the stress-energy tensor's trace, hence $P = (1/3) \sum_{i=1}^{3} T^{ii}$.

   (a) Determine $T^{\mu\nu}$ for a single particle of mass m with trajectory $\vec{x}_0(t)$ and momentum $\vec{p}_0(t)$. Compare to the energy momentum tensor of an ideal fluid.

   (b) Determine $T^{\mu\nu}$ for a photon of frequency ω with trajectory $\vec{x}_0(t)$.

   (c) How is the energy density related to the pressure in the two cases discussed?

2. The hydrodynamical equations describing mass conservation, momentum conservation, and energy conservation for an ideal fluid are

$$\partial_t \rho + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \,, \tag{3.68}$$

$$\partial_t \vec{v} + \left(\vec{v} \cdot \vec{\nabla}\right) \vec{v} = -\frac{\vec{\nabla} P}{\rho} \,, \tag{3.69}$$

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) = -P\vec{\nabla} \cdot \vec{v} \,. \tag{3.70}$$

respectively.

(a) Show using equation (3.70) that an isothermal ideal fluid, i.e. a fluid with constant temperature $T(x, t) = T_0$, is also incompressible, $\vec{\nabla} \cdot \vec{v} = 0$.

(b) Show that for a spherically symmetric and isothermal flow of an ideal gas, equations (3.68) through (3.70) simplify to

$$\partial_t \rho + v \partial_r \rho = 0 \;, \quad \partial_t v - \frac{2v^2}{r} = -c_s^2 \partial_r \ln \rho \;, \tag{3.71}$$

where $c_s \equiv k_B T_0 / m$ is a characteristic thermal speed.

3.2 Relativistic Hydrodynamics

This section is a detour from the main track of this book in so far as General Relativity is otherwise avoided. Yet, it is an irresistible temptation to show how generally-relativistic, ideal hydrodynamics emerges simply if the partial derivatives in the covariant conservation equations (3.33) are replaced by covariant derivatives, and Poisson's equation by (the appropriate limit of) Einstein's field equation. The first main result are the relativistic versions (3.81) and (3.82) of the continuity and Euler equations. In the limit of weak gravitational fields, the relativistic generalisations (3.95) of these equations are derived. Together with the gravitational field equation in the same limit, the final set of hydrodynamical equations is given by (3.106). Perturbative analysis then yields the linear, second-order evolution equation (3.116) for the fluid density.

3.2.1 Hydrodynamic Equations

We shall now derive the ideal hydrodynamic equations from the generally-relativistic equation of local energy conservation. We do this for one specific reason. In the preceding section, we have derived the equations of ideal hydrodynamics by taking appropriate moments of the Boltzmann equation. In that derivation, it has become clear how ideal hydrodynamics builds upon the fluid approximation, and how viscosity and other transport processes such as heat conduction arise if the ideal-fluid approximation is gradually released. Yet, that derivation does not easily allow incorporating the main repercussions of General Relativity in hydrodynamics, which arise because pressure has inertia and contributes as a source to the gravitational field. Therefore, we give this relativistic derivation of the hydrodynamical equations here, borrowing from the differential-geometric formalism of General Relativity without detailed explanation, and contrasting the generally-relativistic hydrodynamic equations at the end with their Newtonian analoga. Our main motivation is that sometimes fluids occur in astrophysics which either move relativistically or whose pressure is comparable to their energy density. In both cases, the classical Newtonian hydrodynamical equations are suspect, and their relativistic counterparts should be used instead.

Readers unfamiliar with general relativity might wish to skip the following subsections, returning when the equations of relativistic hydrodynamics will be summarised and compared to the Newtonian equations.

We begin with the equation of local energy-momentum conservation,

$$\nabla_\nu T^{\mu\nu} = 0 \,, \tag{3.72}$$

which states that the covariant four-divergence of the energy-momentum tensor T has to vanish. This is an immediate consequence of Einstein's field equations. By the second contracted Bianchi identity, the covariant divergence of the Einstein tensor G vanishes identically, so the covariant divergence of the energy-momentum tensor needs to vanish as well.

At this level, we only need to specify that the covariant derivative ∇ is a bi-linear map of (tangent) vectors $(x, y) \in TM$ to a manifold M into the real numbers,

$$\nabla : TM \times TM \to \mathbb{R} \,, \quad (x, y) \mapsto \nabla_x y \,, \tag{3.73}$$

satisfying the Leibniz (product) rule,

$$\nabla_x (fy) = \mathrm{d}f(x)y + f\nabla_x y \,, \tag{3.74}$$

with functions f.

Caution The connection conventionally used in general relativity is specified by two further conditions: it is supposed to be symmetric (torsion-free) and metric-compatible ($\nabla g = 0$). ◀

In a coordinate basis of tangent space, the covariant derivatives are uniquely represented by the Christoffel symbols. More generally, in an arbitrary basis $\{e_\mu\}$ of tangent space, the covariant derivative is defined by the connection 1-forms,

$$\nabla_x e_\mu = \omega^\nu{}_\mu (x) e_\nu \,. \tag{3.75}$$

We now choose to insert the energy-momentum tensor of an ideal fluid,

$$T = \left(\rho c^2 + p\right) u \otimes u - pg \,, \tag{3.76}$$

which is spanned by the only two tensors available in relativistically flowing ideal fluid, namely the tensor product of the four-velocity u with itself and the metric tensor g. The local fluid properties are given by the density ρ and the pressure p measured by the observer flowing with the four-velocity u. Writing the energy-momentum tensor as in (3.76) implies that the four-velocity u must be dimension-less, and thus be measured in units of the light speed c. The components of the energy-momentum tensor T, without specifying the basis vectors yet, are

$$T^{\mu\nu} = \left(\rho c^2 + p\right) u^\mu u^\nu - pg^{\mu\nu} \,. \tag{3.77}$$

Inserting these into the local conservation equation (3.72) gives

$$u^\mu \nabla_u \left(\rho c^2 + p\right) + u^\mu \left(\rho c^2 + p\right) \nabla \cdot u + \left(\rho c^2 + p\right) \nabla_u u^\mu + \nabla^\mu p = 0 \tag{3.78}$$

if we specify the covariant derivative ∇ as usual to be metric, requiring $\nabla g = 0$.

We now project equation (3.78) first on the local time direction by contracting it with the (dual) four-velocity u_μ, and then on the three-space perpendicular to the four-velocity. By their construction, these projections will yield the time and space components of the local conservation equation (3.72), which generalise the continuity and Euler equations.

By definition of the proper time τ, the four-velocity must be normalised by

$$\langle u, u \rangle = u_\mu u^\mu = -1 \,. \tag{3.79}$$

In particular, this normalisation condition implies that

$$0 = \nabla_u \left(u_\mu u^\mu\right) = 2u_\mu \nabla_u u^\mu \ . \tag{3.80}$$

Taking (3.79) and (3.80) into account, contracting (3.78) with u_μ gives the relativistic continuity equation

$$\nabla_u \left(\rho c^2\right) + \left(\rho c^2 + p\right) \nabla \cdot u = 0 \ , \tag{3.81}$$

while its spatial projection by contraction with the projection tensor $\pi_{\alpha\mu} = g_{\alpha\mu} + u_\alpha u_\mu$ yields the relativistic Euler equation

$$\left(\rho c^2 + p\right) \nabla_u u_\alpha + \nabla_\alpha p + u_\alpha \nabla_u p = 0 \ . \tag{3.82}$$

It can easily be seen that $\pi_{\alpha\mu}$ is a projection tensor perpendicular to the four velocity since it maps the four-velocity to zero,

$$\pi_{\alpha\mu} u^\mu = \left(g_{\alpha\mu} + u_\alpha u_\mu\right) u^\mu = u_\alpha - u_\alpha = 0 \ . \tag{3.83}$$

?

Projection tensors π (or, more generally, projections) need to be idempotent, i.e. they need to satisfy $\pi^2 = \pi$. Why is this so? Show that $\pi = g + u \otimes u$ is indeed idempotent. Written in terms of tensor components, show that ${\pi_\alpha}^\beta \pi_{\beta\mu} = \pi_{\alpha\mu}$.

Equations (3.81) and (3.82) form the basis for the following calculations. What do they mean?

The continuity equation (3.81) begins with the covariant derivative of ρc^2 in the direction of the local four-velocity. This generalises the time derivative of the matter density ρ in the continuity equation in three ways. First, the derivative with respect to the coordinate time t is replaced by a derivative with respect to proper time; second, the partial derivative is replaced by a covariant derivative; and third, the matter density is replaced by the energy density ρc^2. The second term in the continuity equation generalises the divergence of the velocity field to the four-divergence of the four-velocity, multiplied with the energy density plus the pressure rather than the density alone: The relativistic continuity equation automatically contains the contribution of pressure-volume work to energy conservation.

The Euler equation starts with the four-acceleration, i.e. the covariant derivative of the four-velocity into the direction of the local four velocity itself. The prefactor $(\rho c^2 + p)$ shows the inertia of pressure. The second term is the pressure gradient, while the third term adds a proper time derivative of the pressure times the flow velocity.

3.2.2 Hydrodynamics in a Weak Gravitational Field

We now proceed to specialise the generally-relativistic continuity and Euler equations, (3.81) and (3.82), to weak gravitational fields. In any metric theory of gravity, in the weak-field limit, the line element can be expressed by means of the two Bardeen potentials ϕ, ψ as

$$\mathrm{d}s^2 = -(1 + 2\phi)\, c^2 \mathrm{d}t^2 + (1 + 2\psi)\, \mathrm{d}\vec{x}^2 \ . \tag{3.84}$$

Both potentials are given in units of c^2, thus dimension-less, and they are assumed to be small, $\phi, \psi \ll 1$. For simplicity, we further take the potentials to be time-independent, $\dot{\phi} = 0 = \dot{\psi}$. The line element (3.84) suggests introducing the dual basis

$$\theta^0 = (1+\phi)c\mathrm{d}t\ , \quad \theta^i = (1+\psi)\mathrm{d}x^i \tag{3.85}$$

and its orthonormal basis

$$e_0 = (1-\phi)c^{-1}\partial_t\ , \quad e_i = (1-\psi)\partial_i\ . \tag{3.86}$$

By this choice of the (dual) basis, the components of the metric become Minkowskian, $g_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$. By means of Cartan's first structure equation, the dual basis $\{\theta^\mu\}$ implies the connection forms

$$\omega^0{}_i = \phi_i\theta^0\ , \quad \omega^i{}_j = \psi_j\theta^i - \psi_i\theta^j\ . \tag{3.87}$$

Now, in a coordinate basis, the four-velocity is

$$u = \tilde{u}^\mu\partial_\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\partial_\mu\ . \tag{3.88}$$

From the line element (3.84), we can read off the proper-time element

$$\mathrm{d}\tau = \left[(1+2\phi) - (1+2\psi)\vec{\beta}^2\right]^{1/2} c\mathrm{d}t \approx \left(1+\phi-\frac{\vec{\beta}^2}{2}\right)c\mathrm{d}t\ , \tag{3.89}$$

valid to first and relevant order in ϕ and $\vec{\beta}^2$. Here, as usual, $\vec{\beta} = \vec{v}/c = \dot{\vec{x}}/c$, and the dot abbreviates the derivative with respect to the *coordinate* time. Thus, to the same order in ϕ and β, the four-velocity is

$$u = \left(1-\phi+\frac{\vec{\beta}^2}{2}\right)\frac{\mathrm{d}x^\mu}{c\mathrm{d}t}\partial_\mu = \left(1-\phi+\frac{\vec{\beta}^2}{2}\right)\left(\partial_0 + \beta^i\partial_i\right) \tag{3.90}$$

in the coordinate basis. Its components in the basis $\{e_\mu\}$ introduced in (3.85) are then determined by $u^\mu = \theta^\mu(u)$ or, again to first order in ϕ and $\vec{v}^2$,

$$u^0 = 1 + \frac{\beta^2}{2}\ , \quad u^i = \beta^i\ . \tag{3.91}$$

Using now the expressions

$$\nabla_\mu f = e_\mu f\ , \quad \nabla_u f = u^\mu\nabla_\mu f = \left(1-\phi+\frac{\beta^2}{2}\right)c^{-1}\dot{f} + \beta^i\partial_i f \tag{3.92}$$

for arbitrary scalar functions f and

$$\nabla_v u^\mu = \mathrm{d}u^\mu(v) + u^\nu\omega^\mu{}_\nu(v) \tag{3.93}$$

for the component μ of the covariant derivative of a vector u into the direction v, we can finally bring the hydrodynamic equations (3.81) and (3.82) into the form

$$\begin{aligned}
\left(1-\phi+\frac{\beta^2}{2}\right)\dot{\rho}c + \left(\vec{\beta}\cdot\vec{\nabla}\right)\rho c^2 + \left(\rho c^2 + p\right)\left[\vec{\nabla}\cdot\vec{\beta} + c^{-1}\partial_t\left(\frac{\beta^2}{2}\right)\right] &= 0\ , \\
\left(\rho c^2 + p\right)\partial_t\left(\frac{\beta^2}{2}\right) + \beta^2\dot{p} + c\vec{\beta}\cdot\vec{\nabla}p &= 0\ , \\
\left(\rho c^2 + p\right)\left[c^{-1}\dot{\vec{\beta}} + \left(\vec{\beta}\cdot\vec{\nabla}\right)\vec{\beta} + \vec{\nabla}\phi\right] + (1-\psi)\vec{\nabla}p + c^{-1}\dot{p}\vec{\beta} + \vec{\beta}\left(\vec{\beta}\cdot\vec{\nabla}\right)p &= 0\ .
\end{aligned} \tag{3.94}$$

In these equations, $\vec{\nabla}$ with a vector arrow is now specialised to be the ordinary gradient operator in three-dimensional, Euclidean space. The second of these equations, which is the time component of the Euler equation, shows that the term $\vec{\beta} \cdot \vec{\nabla} p$ is of order β^2, thus $\psi \vec{\nabla} p$ is of order $\beta^3 \approx 0$ because the potential ψ is itself of order β^2. The continuity equation, to linear order in β, and the Euler equation, to quadratic order in v, are thus

$$\dot{\rho} c + \left(\vec{\beta} \cdot \vec{\nabla}\right) \rho c^2 + \left(\rho c^2 + p\right) \vec{\nabla} \cdot \vec{\beta} = 0 \ ,$$
$$\left(\rho c^2 + p\right)\left[c^{-1}\dot{\vec{\beta}} + \left(\vec{\beta} \cdot \vec{\nabla}\right)\vec{\beta} + \vec{\nabla}\phi\right] + \vec{\nabla} p + c^{-1}\dot{p}\vec{\beta} = 0 \ . \tag{3.95}$$

Notice that, reassuringly, all terms in both these equations have the dimension [energy density]/[length].

3.2.3 Gravitational Field Equation

To linear order in ϕ and ψ, the curvature 2-forms implied by the connection 1-forms (3.79) through Cartan's second structure equation are

$$\Omega^0{}_i = \phi_{ij}\theta^j \wedge \theta^0 \ , \quad \Omega^i{}_j = \psi_{jk}\theta^k \wedge \theta^i - \psi_{ik}\theta^k \wedge \theta^j \ . \tag{3.96}$$

From them, the components of the Ricci tensor can be found via

$$R_{\mu\nu} = \Omega^\alpha{}_\mu \left(e_\alpha, e_\nu\right) \ . \tag{3.97}$$

With (3.96), they are

$$R_{00} = \vec{\nabla}^2\phi \ , \quad R_{0i} = 0 \ , \quad R_{ij} = -(\phi + \psi)_{ij} - \delta_{ij}\vec{\nabla}^2\psi \ . \tag{3.98}$$

The Ricci scalar is

$$R = R^\mu{}_\mu = -2\vec{\nabla}^2 \left(\phi + 2\psi\right) \ , \tag{3.99}$$

and thus the components of the Einstein tensor become

$$G_{00} = -2\vec{\nabla}^2\psi \ , \quad G_{0i} = 0 \ , \quad G_{ij} = -(\phi + \psi)_{ij} + \delta_{ij}\vec{\nabla}^2(\phi + \psi) \ . \tag{3.100}$$

?

Beginning from the components (3.98) of the Ricci tensor, confirm by your own calculation that the Einstein tensor has the components (3.100).

With (3.100), the time-time component of the field equations gives

$$-\vec{\nabla}^2\psi = \frac{4\pi G}{c^4}\left[\rho c^2 + \beta^2\left(\rho c^2 + p\right)\right] \ , \tag{3.101}$$

while the spatial trace of the field equations yields

$$\vec{\nabla}^2(\phi + \psi) = \frac{4\pi G}{c^4}\left[3p + \beta^2\left(\rho c^2 + p\right)\right] \ . \tag{3.102}$$

The sum of the latter two equations gives the generalised Poisson equation

$$\vec{\nabla}^2\phi = \frac{4\pi G}{c^4}\left[\rho c^2 + 3p + 2\beta^2\left(\rho c^2 + p\right)\right] \ . \tag{3.103}$$

The trace of the field equations is

$$\vec{\nabla}^2\phi + 2\vec{\nabla}^2\psi = \frac{4\pi G}{c^4}\left(3p - \rho c^2\right) \ , \tag{3.104}$$

and their off-diagonal components require

$$-(\phi + \psi)_{ij} = \frac{8\pi G}{c^4}\left(\rho c^2 + p\right)\beta_i\beta_j \ . \tag{3.105}$$

3.2.4 The Combined Set of Equations

Thus, to lowest relevant order in ϕ, ψ and $\beta^2 = v^2/c^2$, the combined hydrodynamic and gravitational equations are

$$\dot{\rho} + \left(\vec{v}\cdot\vec{\nabla}\right)\rho + \left(\rho + \frac{p}{c^2}\right)\vec{\nabla}\cdot\vec{v} = 0 \;,$$
$$\dot{\vec{v}} + \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v} = -\vec{\nabla}\Phi - \frac{\vec{\nabla}p + \dot{p}\vec{v}/c^2}{\rho + p/c^2} \;,$$
$$\vec{\nabla}^2\Phi = 4\pi G\left(\rho + \frac{3p}{c^2}\right) \;. \tag{3.106}$$

They generalise the Newtonian hydrodynamic equations

$$\dot{\rho} + \left(\vec{v}\cdot\vec{\nabla}\right)\rho + \rho\vec{\nabla}\cdot\vec{v} = 0 \;,$$
$$\dot{\vec{v}} + \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v} = -\vec{\nabla}\Phi - \frac{\vec{\nabla}p}{\rho} \;,$$
$$\vec{\nabla}^2\Phi = 4\pi G\rho \;, \tag{3.107}$$

where $\Phi = c^2\phi$ is the Newtonian gravitational potential in physical units. Comparing (3.106) and (3.107), one clearly sees the pressure-volume work in the continuity equation, the inertia of the mass-density equivalent p/c^2 of the pressure in the Euler equation and the contribution of $3p/c^2$ to the source of the gravitational field. Notice also the additional force term $\propto \dot{p}\vec{v}/c^2$ in the Euler equation.

3.2.5 Perturbative Analysis

Let us now continue with a perturbative analysis of the set of equations (3.106). As usual, we assume that a smooth background solution is already given, which is indicated by a subscript 0. We thus have a set of fields $(\rho_0, p_0, \vec{v}_0, \phi_0)$ which separately satisfy Eqs. (3.106). They are perturbed by small deviations $(\delta\rho, \delta p, \delta\vec{v}, \delta\phi)$. The equations will be linearised in these perturbations, meaning that terms will be dropped that are of quadratic or higher order in the perturbations.

We transform into a coordinate system comoving with the unperturbed flow, which allows us to set $\vec{v}_0 = 0$. We assume that the perturbations are small compared to the overall length scale of the unperturbed solution, hence gradients of the background solution can be neglected. Finally, we assume that the fluid has a polytropic equation of state,

$$p = \bar{p}\left(\frac{\rho}{\bar{\rho}}\right)^{\gamma} \;, \tag{3.108}$$

where $(\bar{\rho}, \bar{p})$ are arbitrary reference values for the density and the pressure and γ is the adiabatic index of the fluid. Since the squared sound speed is

$$c_\mathrm{s}^2 = \frac{\partial p}{\partial \rho} \tag{3.109}$$

at constant entropy, the pressure fluctuations can be written as $\delta p = c_\mathrm{s}^2 \delta\rho$. We express the density fluctuation by the dimension-less density contrast

$$\delta = \frac{\delta\rho}{\rho_0}\,, \quad \delta\rho = \rho_0\delta\,, \tag{3.110}$$

which allows to write the pressure fluctuation as

$$\delta p = c_\mathrm{s}^2 \rho_0 \delta\,. \tag{3.111}$$

Substituting $\rho = \rho_0 + \delta\rho = \rho_0(1+\delta)$ and $\vec{v} = \vec{v}_0 + \delta\vec{v} = \delta\vec{v}$ into the continuity equation gives, to lowest order in the perturbations,

$$\dot{\rho}_0 = 0\,, \tag{3.112}$$

and to first order

$$\dot{\delta} + \left(1 + \frac{p_0}{\rho_0 c^2}\right)\vec{\nabla}\cdot\delta\vec{v} = 0\,. \tag{3.113}$$

By the polytropic equation-of-state, (3.112) also implies $\dot{p}_0 = 0$. Then, to first order in the perturbations, Euler's equation and the generalised Poisson equation are reduced to

$$\begin{aligned} \delta\dot{\vec{v}} &= -\vec{\nabla}\delta\Phi - \frac{c_\mathrm{s}^2 \vec{\nabla}\delta}{1 + \frac{p_0}{\rho_0 c^2}}\,, \\ \vec{\nabla}^2 \delta\Phi &= 4\pi G \rho_0 \delta \left(1 + \frac{3c_\mathrm{s}^2}{c^2}\right)\,. \end{aligned} \tag{3.114}$$

Taking the time derivative of the continuity equation (3.113) and the divergence of the Euler equation from (3.114) transforms these equations into

$$\begin{aligned} \ddot{\delta} + \left(1 + \frac{p_0}{\rho_0 c^2}\right)\vec{\nabla}\cdot\delta\dot{\vec{v}} &= 0\,, \\ \vec{\nabla}\cdot\delta\dot{\vec{v}} &= -\vec{\nabla}^2\delta\Phi - \frac{c_\mathrm{s}^2 \vec{\nabla}^2\delta}{1 + \frac{p_0}{\rho_0 c^2}}\,. \end{aligned} \tag{3.115}$$

Eliminating the divergence of the peculiar acceleration, $\vec{\nabla}\cdot\delta\dot{\vec{v}}$, between these equations and inserting the generalised Poisson equation from (3.114) then leads to the evolution equation

$$\ddot{\delta} - 4\pi G \rho_0 \delta \left(1 + \frac{3c_\mathrm{s}^2}{c^2}\right)\left(1 + \frac{p_0}{\rho_0 c^2}\right) - c_\mathrm{s}^2 \vec{\nabla}^2 \delta = 0 \tag{3.116}$$

for the density contrast δ. In the non-relativistic limit, when the sound speed c_s is small compared to the light speed c and the pressure p_0 is negligible compared to the rest-energy density $\rho_0 c^2$, this linear evolution equation for the density contrast shrinks to

$$\ddot{\delta} - 4\pi G \rho_0 \delta - c_\mathrm{s}^2 \vec{\nabla}^2 \delta = 0\,. \tag{3.117}$$

Example: Without gravity

Some special cases should now be instructive and illustrate why we went through this analysis here. Let us first ignore gravity completely. Then, the second term in (3.116) disappears altogether because it originates from gravity alone. Ignoring gravity can formally be expressed by setting the Newtonian gravitational constant to zero, $G = 0$, and thus suppress all gravitational coupling. Then, the density fluctuations δ are found to obey the wave equation

$$\Box\delta = 0 \,, \tag{3.118}$$

where the sound speed c_s appears as the characteristic velocity in the d'Alembert operator. The density contrast then undergoes ordinary sound waves. ◂

Example: With gravity on a non-relativistic background

If gravity is switched back on, but the background remains non-relativistic, Eq. (3.116) simplifies to

$$\ddot{\delta} - 4\pi G\rho_0\delta - c_\mathrm{s}^2\vec{\nabla}^2\delta = 0 \,. \tag{3.119}$$

If we expand δ into plane waves, the Laplacian is replaced by the negative square of the wave number k, and δ obeys

$$\ddot{\delta} - \left(4\pi G\rho_0 - c_\mathrm{s}^2k^2\right)\delta = 0 \,. \tag{3.120}$$

This is an ordinary oscillator equation, with

$$c_\mathrm{s}^2k^2 - 4\pi G\rho_0 = \omega^2 \tag{3.121}$$

taking the role of the squared frequency. If $\omega^2 > 0$, i.e. if k is larger than the so-called Jeans wave number

$$k_\mathrm{J} = \left(\frac{4\pi G\rho_0}{c_\mathrm{s}^2}\right)^{1/2} \,, \tag{3.122}$$

the solutions oscillate like sound waves, satisfying the dispersion relation

$$\omega = c_\mathrm{s}\sqrt{k^2 - k_\mathrm{J}^2} \,. \tag{3.123}$$

Otherwise, if k is smaller than the Jeans wave number, there is an exponentially growing and an exponentially decaying mode of the density fluctuations. ◂

Example: With gravity on a relativistic background

If, finally, the background fluid is relativistic, we have the full equation

$$\ddot{\delta} - \left[4\pi G\rho_0\left(1+\frac{3c_s^2}{c^2}\right)\left(1+\frac{p_0}{\rho_0 c^2}\right) - c_s^2k^2\right]\delta = 0 \tag{3.124}$$

for plane waves of wave number k. Suppose, for example, we have a plasma tightly coupled to a dominating photon gas like in the early universe. Then, the fluid is relativistic, $p_0 \approx \rho_0 c^2/3$ and $c_s^2 \approx c^2/3$, and

$$\left(1+\frac{3c_s^2}{c^2}\right)\left(1+\frac{p_0}{\rho_0 c^2}\right) \approx \frac{8}{3}\ . \tag{3.125}$$

The Jeans wave number then changes to

$$k_J = \left(\frac{32\pi G\rho_0}{3c_s^2}\right)^{1/2} = \left(\frac{32\pi G\rho_0}{c^2}\right)^{1/2}\ , \tag{3.126}$$

which is typically much smaller than for a non-relativistic fluid. Acoustically oscillating perturbations are thus possible in a much wider range of scales in a relativistic than in a non-relativistic fluid, and growth or decay of perturbations is possible only for very large perturbations. ◂

3.3 Viscous hydrodynamics

So far, we have considered ideal fluids, whose particles have a negligibly small mean free path. In this section, we shall loosen this approximation and allow a very small, but finite mean free path. The fluid particles can now move relative to the mean flow and transport fluid properties by small distances, in particular mass, momentum and energy. The transport of momentum causes friction and energy dissipation, the transport of energy gives rise to heat conduction. The first important result is the diffusive extension of the energy-momentum tensor (3.143) which can then be introduced into the conservation equation to derive the Navier-Stokes equation (3.148) and the energy-conservation equation (3.155) containing heat flow and dissipation. Finally, we introduce gravitational forces into the equations of viscous hydrodynamics and derive the tensor virial theorem (3.189).

3.3.1 Diffusion of particles, momentum and internal energy

Previously, we have assumed that our fluid is ideal, that is, that the mean-free path λ is negligibly small. We have used this implicitly when we set the momentum-space integrals over the collision term to zero. If we cannot neglect the mean-free path any more, we must take into account that particles may move over small, but non-vanishing distances and thereby carry their physical properties with them. In that way, transport phenomena occur over small distance scales.

Let us begin with a simple example. Suppose there is a homogeneous ideal fluid, into which we place a screen of the small cross-sectional area dA. For definiteness, we set up a coordinate system such that the screen is perpendicular to one of the coordinate axes, say the x axis, which it may intersect at the coordinate origin. This screen and the coordinate system may flow with the mean fluid velocity.

Particles will move by random motion from one side of the screen to the other. Let $n(0)$ be their mean number density at the location of the screen, and the screen be small enough for us to neglect any change of the number density across the screen. If the particle number density behind the screen is the same as in front of the screen, the same number of particles will cross the screen per unit time in the positive as in the negative x direction, and the net number of particles flowing through the screen will be zero.

Now let us gradually relax this stationary situation by imagining a number-density gradient along the x direction (cf. Fig. 3.3). Then, there will be fewer particles behind than before the screen, and even though their random velocities in the $\pm x$ directions will be the same, more particles will flow down than up the gradient. Let $\bar{u}$ be a characteristic velocity of the particles. Since their random velocities $\vec{u}$ will average to zero, $\bar{u}$ could be set to the root mean-square velocity, $\bar{u} = \langle \vec{u}^2 \rangle^{1/2}$. How exactly $\bar{u}$ and $\langle \vec{u}^2 \rangle^{1/2}$ relate depends on the velocity distribution of the particles, which is however irrelevant for our purposes.

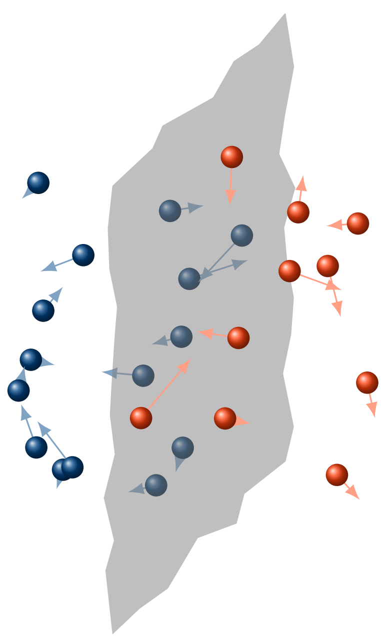

Figure 3.3 Particle diffusion: If there are more particles on one side of the imagined screen than on the other, such as there are more blue than red particles in this example, particles will effectively diffuse from the denser region to the less dense.

Then if the particle velocities are randomly oriented, the velocity in the x direction, perpendicular to the screen, will be of order $\bar{u}/\sqrt{3}$. Since approximately half of the particles will move into the positive x direction, the number of particles N moving through the screen in either direction per unit time is

$$\frac{\mathrm{d}N}{\mathrm{d}t} \approx -\frac{\bar{u}}{2\sqrt{3}}\mathrm{d}A\,[n(x+\lambda) - n(x-\lambda)] \ , \tag{3.127}$$

where λ is the mean free path of the particles. If λ is finite, but small, we can replace the difference in particle number densities by a derivative to find the particle current density

$$j_\mathrm{p} = \frac{\mathrm{d}N}{\mathrm{d}A\mathrm{d}t} \approx -\frac{\bar{u}\lambda}{\sqrt{3}}\frac{\partial n}{\partial x} \ . \tag{3.128}$$

In three dimensions, the derivative with respect to x is replaced by the gradient,

$$\vec{j}_\mathrm{p} = -\frac{\bar{u}\lambda}{\sqrt{3}}\vec{\nabla}n \ . \tag{3.129}$$

Gradients in particle number densities drive particle diffusion. Inserting this current together with the particle number density into the continuity equation for the particles gives Fick's (second) law for diffusive particle transport,

$$\partial_t n + \vec{\nabla}\cdot\vec{j}_\mathrm{p} = 0 \quad \Rightarrow \quad \partial_t n = \vec{\nabla}\cdot\left(D\vec{\nabla}n\right) \ , \quad D = \frac{\bar{u}\lambda}{\sqrt{3}} \ . \tag{3.130}$$

Recall that the expression given here for the diffusion coefficient D has been heuristically derived. More precise definitions can be given if the probability distribution of the random velocities is known.

Let us now apply the same approach to momentum and energy transport. Consider how particles transport a velocity component v^i diffusively into the x direction. If v^i changes with x, the velocity component v^i of the particles diffusing towards the positive x direction differs from the v^i that the particles transport towards the negative x direction. By essentially the same argument that led to (3.127), we find the current density component $(j_\mathrm{v})^i_x$ of v^i

$$(j_\mathrm{v})^i_x = -\frac{n\bar{u}\lambda}{\sqrt{3}}\frac{\partial v^i}{\partial x} \ . \tag{3.131}$$

?

Derive (3.131) in a way similar to the derivation of (3.128).

The diffusive transport of the velocity component v^i into the spatial direction x_j can accordingly be described by the rank-2 tensor

$$(j_\mathrm{v})^i{}_j = -\frac{n\bar{u}\lambda}{\sqrt{3}}\frac{\partial v^i}{\partial x^j} \ . \tag{3.132}$$

This suggests that the stress-energy tensor $\bar{T}_\mathrm{d}$ describing diffusive momentum transport should be proportional to the tensor of spatial velocity derivatives,

$$\bar{T}_\mathrm{d} \propto -\left(\vec{\nabla}\otimes\vec{v}\right)^\top \ , \tag{3.133}$$

with a proportionality constant giving the right-hand side the appropriate dimension of a momentum current density.

Energy transport by diffusion is easily completed. Completely analogously to the previous derivations, we find the diffusive current density of the internal energy

$$\vec{q}_\varepsilon = -\frac{n\bar{u}\lambda}{\sqrt{3}}\vec{\nabla}\varepsilon \ . \tag{3.134}$$

We can express the gradient of the internal energy by a temperature gradient and obtain

$$\vec{q}_\varepsilon = -\frac{n\bar{u}\lambda}{\sqrt{3}}\frac{\mathrm{d}\varepsilon}{\mathrm{d}T}\vec{\nabla}T = -\frac{n\bar{u}\lambda c_\mathrm{v}}{\sqrt{3}}\vec{\nabla}T = -\kappa\vec{\nabla}T\;. \tag{3.135}$$

For the second equality, we have inserted the heat capacity c_v at constant volume, and the last equality defines the heat conductivity κ.

The diffusive stress-energy tensor requires further consideration. While the velocity-gradient tensor $\vec{\nabla}\otimes\vec{v}$ may be asymmetric, the stress-energy tensor should be symmetric. This suggests to assume that the diffusive stress-energy tensor should be set proportional to the symmetric part of $\vec{\nabla}\otimes\vec{v}$, or

$$\bar{T}_\mathrm{d} \propto -\left[\left(\vec{\nabla}\otimes\vec{v}\right) - \left(\vec{\nabla}\otimes\vec{v}\right)^\top\right]\;. \tag{3.136}$$

This is reasonable also because of the following consideration. If a system of particles rotates like a solid body of angular velocity $\vec{\omega}$, i.e. with the velocity field

$$\vec{v} = \vec{\omega}\times\vec{r}\;,\quad v^j = \varepsilon^j{}_{kl}\,\omega^k x^l\;, \tag{3.137}$$

no momentum transport should occur. The derivatives of the velocity components (3.137) are

$$\partial_i v^j = \partial_i\left(\varepsilon^j{}_{kl}\,\omega^k x^l\right) = \varepsilon^j{}_{kl}\,\omega^k\delta^l_i = \varepsilon^j{}_{ki}\,\omega^k\;, \tag{3.138}$$

which is manifestly antisymmetric because of the antisymmetry of the Levi-Civita symbol. Excluding momentum-transport effects in systems rotating like solid bodies thus also argues for setting the diffusive stress-energy tensor proportional to the symmetrised velocity-gradient tensor.

?

Explain the factor of $2/3$ in the trace-free part (3.141) of the velocity-gradient tensor.

It is further often convenient to distinguish between divergent or convergent flows, for which

$$\vec{\nabla}\cdot\vec{v} = \partial_i v^i = \mathrm{Tr}\left(\vec{\nabla}\otimes\vec{v}\right) \neq 0\;, \tag{3.139}$$

and so-called shear flows, for which the trace vanishes,

$$\mathrm{Tr}\left(\vec{\nabla}\otimes\vec{v}\right) = 0\;. \tag{3.140}$$

We thus split up the symmetrised velocity-gradient tensor into a trace-free part

$$\left(\vec{\nabla}\otimes\vec{v}\right) + \left(\vec{\nabla}\otimes\vec{v}\right)^\top + \frac{2}{3}\vec{\nabla}\cdot\vec{v}\,\mathbb{1}_3 \tag{3.141}$$

and a diagonal part carrying the trace,

$$\vec{\nabla}\cdot\vec{v}\,\mathbb{1}_3\;, \tag{3.142}$$

and assemble the diffusive stress-energy tensor from these two contributions separately,

$$\bar{T}_\mathrm{d} = -\eta\left[\left(\vec{\nabla}\otimes\vec{v}\right) + \left(\vec{\nabla}\otimes\vec{v}\right)^\top - \frac{2}{3}\vec{\nabla}\cdot\vec{v}\,\mathbb{1}_3\right] - \zeta\vec{\nabla}\cdot\vec{v}\,\mathbb{1}_3\;. \tag{3.143}$$

The two constants η and ζ introduced here represent the viscosity of the fluid.

The components of the stress-energy tensor must have the dimension of a momentum current density, hence

$$\left[T_{\rm d}^{ij}\right] = \frac{\rm g\,cm}{\rm s}\frac{1}{\rm cm^2\,s} = \frac{\rm g}{\rm cm\,s^2}\ . \tag{3.144}$$

Since the velocity gradient components have the dimension $\rm s^{-1}$, the viscosity constants must have the dimension

$$[\eta] = \frac{\rm g}{\rm cm\,s} = [\zeta]\ . \tag{3.145}$$

3.3.2 The equations of viscous hydrodynamics

The diffusion of fluid particles cannot affect mass conservation, so the continuity equation (3.36) for the mass density must remain unchanged. However, the preceding considerations of diffusive particle, energy and momentum transport have shown that we have to augment the stress-energy tensor of an ideal fluid by the diffusive stress-energy tensor,

$$\bar{T} \to \bar{T} + \bar{T}_{\rm d}\ . \tag{3.146}$$

Since momentum conservation is expressed by the spatial components $\partial_\mu T^{\mu i} = 0$ of the conservation equation $\partial_\mu T^{\mu\nu} = 0$, the additional, diffusive part of the stress-energy tensor creates the further terms

$$\begin{aligned} \vec{\nabla}\cdot\bar{T}_{\rm d} &= -\eta\left[\vec{\nabla}^2\vec{v} + \vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right) - \frac{2}{3}\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right)\right] - \zeta\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right) \\ &= -\eta\vec{\nabla}^2\vec{v} - \left(\zeta + \frac{\eta}{3}\right)\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right) \end{aligned} \tag{3.147}$$

in the momentum-conservation equation (3.54). With those terms, it turns into the Navier-Stokes equation

$$\rho\left(\partial_t + \vec{v}\cdot\vec{\nabla}\right)\vec{v} + \vec{\nabla}P = \eta\vec{\nabla}^2\vec{v} + \left(\zeta + \frac{\eta}{3}\right)\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right)\ . \tag{3.148}$$

?

Compare the Navier-Stokes equation (3.148) to the Euler equation (3.54) and discuss (with yourself or somebody else) the physical meaning of the difference between the two.

In the energy-conservation equation, we must first of all take the diffusive transport of the internal energy into account, thus

$$\vec{q} \to \vec{q} + \vec{q}_\varepsilon = \vec{q} - \kappa\vec{\nabla}T \tag{3.149}$$

needs to be substituted in (3.40). However, this is not all, since the diffusive momentum-current density corresponds to a force per unit area, or a pressure. This force, times the flow velocity, is the internal work carried out per unit time by the diffusing particles on the fluid itself; in other words, it is the energy per unit time dissipated by friction. The current density of this friction work is the flow velocity times the force per unit area,

$$\vec{q}_{\rm fr} = -\bar{T}_{\rm d}\vec{v}\ , \tag{3.150}$$

which must also be added to the energy current density $\vec{q}$. Thus,

$$\vec{q} \to \vec{q} - \kappa\vec{\nabla}T - \bar{T}_{\rm d}\vec{v} \tag{3.151}$$

must be replaced in (3.40). Deriving the final form for the energy-conservation equation, we must finally recall that the momentum-conservation equation has also changed. We used it before to bring the energy-conservation equation into the form (3.60), subtracting the momentum-conservation equation, multiplied with the flow velocity, from the energy-conservation equation. We thus have to subtract a further term $(\vec{\nabla} \cdot \bar{T}_\mathrm{d}) \cdot \vec{v}$ from the energy-conservation equation. Summing up, the right-hand side of the energy-conservation equation must now be replaced by the terms

$$\vec{\nabla} \cdot \left(\kappa \vec{\nabla} T + \bar{T}_\mathrm{d} \vec{v}\right) - \left(\vec{\nabla} \cdot \bar{T}_\mathrm{d}\right) \cdot \vec{v} = \vec{\nabla} \cdot \left(\kappa \vec{\nabla} T\right) + \mathrm{Tr}\left[\bar{T}_\mathrm{d}^\top \left(\vec{\nabla} \otimes \vec{v}\right)\right] . \tag{3.152}$$

Since the diffusive stress-energy tensor is symmetric, any antisymmetric part of the velocity-gradient tensor $\vec{\nabla} \otimes \vec{v}$ would be cancelled in its contraction with $\bar{T}_\mathrm{d}$, hence we can just as well write

$$\mathrm{Tr}\left[\bar{T}_\mathrm{d}^\top \left(\vec{\nabla} \otimes \vec{v}\right)\right] = \mathrm{Tr}\left(\bar{T}_\mathrm{d}^\top Dv\right) , \tag{3.153}$$

where Dv abbreviates the symmetrised velocity-gradient tensor,

$$Dv := \frac{1}{2}\left[\left(\vec{\nabla} \otimes \vec{v}\right) + \left(\vec{\nabla} \otimes \vec{v}\right)^\top\right] . \tag{3.154}$$

The energy-conservation equation for a viscous fluid then reads

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) + P \vec{\nabla} \cdot \vec{v} = \vec{\nabla} \cdot \left(\kappa \vec{\nabla} T\right) + \mathrm{Tr}\left(\bar{T}_\mathrm{d}^\top Dv\right) . \tag{3.155}$$

This intuitive equation shows that temperature gradients cause diffusive heat transport, and viscosity creates heat by friction. Together with the unchanged continuity equation (3.36) for the density ρ, the Navier-Stokes equation (3.148) and the energy-conservation equation (3.155) are the fundamental equations for viscous hydrodynamics.

3.3.3 Entropy

It is instructive to translate the energy-conservation equation (3.155) to an equation explicitly containing the fluid entropy. For doing so, we introduce the internal energy and the entropy per unit mass, $\tilde{\varepsilon}$ and $\tilde{s}$, respectively, by defining

$$\varepsilon = \tilde{\varepsilon}\rho , \quad s = \tilde{s}\rho , \tag{3.156}$$

This enables us to bring the left-hand side of (3.155) into the form

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) + P \vec{\nabla} \cdot \vec{v} = \partial_t (\tilde{\varepsilon}\rho) + \vec{\nabla} \cdot (\tilde{\varepsilon}\rho \vec{v}) + P \vec{\nabla} \cdot \vec{v} . \tag{3.157}$$

Subtracting the continuity equation leaves us with

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) + P \vec{\nabla} \cdot \vec{v} = \rho \left(\partial_t + \vec{v} \cdot \vec{\nabla}\right) \tilde{\varepsilon} + P \vec{\nabla} \cdot \vec{v} = \rho \frac{\mathrm{d}\tilde{\varepsilon}}{\mathrm{d}t} + P \vec{\nabla} \cdot \vec{v} . \tag{3.158}$$

The volume per unit mass, $\tilde{V}$, is the reciprocal density, $\tilde{V} = \rho^{-1}$, hence

$$\frac{\mathrm{d}\tilde{V}}{\mathrm{d}t} = -\rho^{-2} \frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho^{-2} \left(\partial_t + \vec{v} \cdot \vec{\nabla}\right) \rho = \rho^{-1} \vec{\nabla} \cdot \vec{v} , \tag{3.159}$$

where we have used the continuity equation once more in the final step. Solving for the velocity divergence,

$$\vec{\nabla} \cdot \vec{v} = \rho \frac{\mathrm{d}\tilde{V}}{\mathrm{d}t} \ , \tag{3.160}$$

we can write (3.158) as

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) + P \vec{\nabla} \cdot \vec{v} = \rho \left(\frac{\mathrm{d}\tilde{\varepsilon}}{\mathrm{d}t} + P \frac{\mathrm{d}\tilde{V}}{\mathrm{d}t} \right) \ . \tag{3.161}$$

?

What is the physical meaning of the expression $\rho T \mathrm{d}\tilde{s}$?

By the first law of thermodynamics, $T\mathrm{d}\tilde{s} = \mathrm{d}\tilde{\varepsilon} + P\mathrm{d}\tilde{V}$, we can finally identify

$$\partial_t \varepsilon + \vec{\nabla} \cdot (\varepsilon \vec{v}) + P \vec{\nabla} \cdot \vec{v} = \rho T \frac{\mathrm{d}\tilde{s}}{\mathrm{d}t} \tag{3.162}$$

and write the energy-conservation equation (3.155) as an equation for the total time derivative of the specific entropy,

$$\rho T \frac{\mathrm{d}\tilde{s}}{\mathrm{d}t} = \vec{\nabla} \cdot \left(\kappa \vec{\nabla} T \right) + \mathrm{Tr} \left(\bar{T}_\mathrm{d}^\top D v \right) \ . \tag{3.163}$$

This shows explicitly how heat conduction and viscous friction change the entropy. In absence of transport processes, $\kappa = 0 = \eta = \zeta$, the specific entropy is conserved. In particular, flows of ideal fluids are isentropic.

3.3.4 Fluids in a gravitational field

From a consistent, generally-relativistic point of view, fluids in a gravitational field should be treated starting from the covariant, local energy-momentum conservation laws (3.33). The covariant derivative would then automatically take care of gravitational forces. Here, in our non-relativistic, Newtonian approach, we have to add gravitational fields by hand to the fluid equations. We shall do so by deriving the energy-momentum tensor of the free gravitational field, whose space-space components can then be added to the stress-energy tensor T^{ij} of the fluid.

In a specially-relativistic theory for a scalar field ϕ characterised by a Lagrange density $\mathcal{L}(\phi, \partial_\mu \phi)$, the energy-momentum tensor $T^\mu{}_\nu$ of the field is given by the Legendre transform

$$T^\mu{}_\nu = \partial_\nu \phi \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \delta^\mu_\nu \mathcal{L} \ ; \tag{3.164}$$

cf. (1.104) and the explanation given there. The Lagrange density

$$\mathcal{L} = \frac{1}{8\pi G} \partial_\mu \phi \partial^\mu \phi + \phi \rho \tag{3.165}$$

serves our purposes because its Euler-Lagrange equation,

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \ , \tag{3.166}$$

reproduces Poisson's equation if the potential ϕ does not change very rapidly with time,

$$-\partial_0 \partial^0 \phi = c^{-2} \partial_t^2 \phi \ll \partial_i \partial^i \phi \ . \tag{3.167}$$

Then, we can approximate the d'Alembert by the Laplace operator and find the familiar field equation

$$\vec{\nabla}^2\phi = 4\pi G\rho \tag{3.168}$$

relating the potential to the density, i.e. the Poisson equation.

We thus take

$$\mathcal{L}_{\text{free}} = \frac{1}{8\pi G}\partial_\mu\phi\partial^\mu\phi \tag{3.169}$$

as the Lagrange density of the free, Newtonian gravitational field and find the energy-momentum tensor

$$\left(T^\mu{}_\nu\right)_{\text{grav}} = \frac{1}{4\pi G}\left(\partial_\nu\phi\partial^\mu\phi - \frac{1}{2}\delta^\mu_\nu\partial_\alpha\phi\partial^\alpha\phi\right) \tag{3.170}$$

for it. Its stress-energy tensor is then

$$\bar{T}_{\text{grav}} = \frac{1}{4\pi G}\left[\vec{\nabla}\phi\otimes\vec{\nabla}\phi - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2\mathbb{1}_3\right], \tag{3.171}$$

again neglecting the time derivative of ϕ compared to its spatial derivatives. This gravitational stress-energy tensor must now be introduced into the equations for momentum and energy conservation.

The momentum-conservation equation must be augmented by the divergence of $\bar{T}_{\text{grav}}$,

$$\vec{\nabla}\cdot\bar{T}_{\text{grav}} = \frac{1}{4\pi G}\left(\vec{\nabla}^2\phi\right)\vec{\nabla}\phi = \rho\vec{\nabla}\phi\,, \tag{3.172}$$

where the Poisson equation (3.168) was used in the last step. With this additional specific force term, the Navier-Stokes equation becomes

$$\rho\left(\partial_t + \vec{v}\cdot\vec{\nabla}\right)\vec{v} + \vec{\nabla}P = -\rho\vec{\nabla}\Phi + \eta\vec{\nabla}^2\vec{v} + \left(\zeta + \frac{\eta}{3}\right)\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right)\,. \tag{3.173}$$

Caution Of course, we could have guessed the additional gravitational-force term $-\rho\vec{\nabla}\Phi$ in the Navier-Stokes immediately since it simply expresses the gravitational-force density. Returning to the stress-energy tensor of the gravitational field and taking its divergence emphasises the common origin of all force terms in the Euler or Navier-Stokes equations. ◄

In most applications, the stress-energy tensor for the free gravitational field is integrated over the entire volume of a body. If the boundary surface of the integration volume is chosen large enough, we can use Gauss' law to add or subtract arbitrary divergences from $\bar{T}_{\text{grav}}$ without changing the volume integral over it. This allows us to modify the expression for the stress-energy tensor to bring it into more familiar forms that can more easily be interpreted. We shall use the sign $\simeq$ here to express that two expressions for T^{ij}_{grav} differ only by a divergence.

Let us begin with the expression (3.171) and write

$$\begin{aligned}\bar{T}_{\text{grav}} &= \frac{\vec{\nabla}\phi\otimes\vec{\nabla}\phi}{4\pi G} - \frac{\mathbb{1}_3}{8\pi G}\vec{\nabla}\cdot\left(\phi\vec{\nabla}\phi\right) + \frac{\mathbb{1}_3}{2}\phi\rho \\ &\simeq \frac{\vec{\nabla}\phi\otimes\vec{\nabla}\phi}{4\pi G} + \frac{\mathbb{1}_3}{2}\phi\rho\,.\end{aligned} \tag{3.174}$$

The trace of the final expression is

$$\begin{aligned}\operatorname{Tr}\bar{T}_{\text{grav}} &= \frac{1}{4\pi G}\left(\vec{\nabla}\phi\right)^2 + \frac{3}{2}\phi\rho = \frac{1}{4\pi G}\vec{\nabla}\cdot\left(\phi\vec{\nabla}\phi\right) - \rho\phi + \frac{3}{2}\rho\phi \\ &\simeq \frac{1}{2}\rho\phi\,.\end{aligned} \tag{3.175}$$

If we rather begin with the expression $\vec{x} \otimes \bar{T}$ for an arbitrary, not necessarily gravitational stress-energy tensor, we can write

$$-\vec{x} \otimes \left(\vec{\nabla} \cdot \bar{T}\right) = -\vec{\nabla} \cdot \left(\vec{x} \otimes \bar{T}\right) + \bar{T} \simeq \bar{T} \ . \tag{3.176}$$

Applying this result to the gravitational stress-energy tensor, and identifying its divergence (3.172), we find

$$\bar{T}_{\rm grav} \simeq -\rho \vec{x} \otimes \vec{\nabla}\phi \ . \tag{3.177}$$

This leads to Chandrasekhar's expression for the gravitational potential energy, which is often used in stellar dynamics,

$$\int \mathrm{d}^3x \, \bar{T}_{\rm grav} =: U = -\int \mathrm{d}^3x \, \rho \vec{x} \otimes \vec{\nabla}\phi \ . \tag{3.178}$$

From our previous result (3.175), we can further infer that the volume integral over the trace of $\bar{T}_{\rm grav}$ is

$$\int \mathrm{d}^3x \ \mathrm{Tr}\, \bar{T}_{\rm grav} = \mathrm{Tr}\, U = \frac{1}{2} \int \mathrm{d}^3x \, \rho\phi \ . \tag{3.179}$$

Comparing (3.178) and (3.179), we find the useful equality

$$\frac{1}{2} \int \mathrm{d}^3x \, \rho\phi = -\int \mathrm{d}^3x \, \rho \vec{x} \cdot \vec{\nabla}\phi \ . \tag{3.180}$$

3.3.5 The tensor virial theorem

We can now derive an important generalisation of the virial theorem from classical mechanics, which is typically derived there for point particles on bounded orbits. We begin with the inertial tensor of a body, defined by

$$I = \int \mathrm{d}^3x \, \rho \vec{x} \otimes \vec{x} \ . \tag{3.181}$$

Integrating over a fixed volume, the position vectors $\vec{x}$ do not depend on time. The total time derivative of I is

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \left(\partial_t + \vec{v} \cdot \vec{\nabla}\right) \int \mathrm{d}^3x \, \rho \vec{x} \otimes \vec{x} = \int \mathrm{d}^3x \, (\partial_t\rho) \vec{x} \otimes \vec{x} \tag{3.182}$$

because the volume integral does not depend on $\vec{x}$. The continuity equation allows us to continue

$$\begin{aligned}\int \mathrm{d}^3x \, (\partial_t\rho) \vec{x} \otimes \vec{x} &= -\int \mathrm{d}^3x \, \vec{\nabla} \cdot (\rho\vec{v}) \, \vec{x} \otimes \vec{x} \\ &= -\int \mathrm{d}^3x \, \vec{\nabla} \cdot (\rho \vec{x} \otimes \vec{x} \otimes \vec{v}) + \int \mathrm{d}^3x \, \rho \, (\vec{x} \otimes \vec{v} + \vec{v} \otimes \vec{x}) \\ &= \int \mathrm{d}^3x \, \rho \, (\vec{x} \otimes \vec{v} + \vec{v} \otimes \vec{x}) \ . \end{aligned} \tag{3.183}$$

?

Can you confirm the calculation shown in (3.183)?

The second absolute time derivative of the inertial tensor is thus

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = \int \mathrm{d}^3x \, \left[\partial_t (\rho\vec{v}) \otimes \vec{x} + \vec{x} \otimes \partial_t (\rho\vec{v})\right] \ . \tag{3.184}$$

Now, we use (3.176) and take advantage of momentum conservation, $\partial_0 T^{0i} + \partial_j T^{ij} = 0$, to replace the divergence of the stress-energy tensor by the time derivative of the energy-current density $T^{0j} = \rho c v^j$,

$$T^{ij} \simeq x^i \partial_0 T^{0j} = x^i \partial_t \left(\rho v^j\right) . \tag{3.185}$$

Symmetrising this expression,

$$\bar{T} = \frac{1}{2} \left[\vec{x} \otimes \partial_t (\rho \vec{v}) + \partial_t (\rho \vec{v}) \otimes \vec{x}\right] , \tag{3.186}$$

and inserting the result into (3.184), we can finally write

$$\frac{1}{2} \frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = \int \mathrm{d}^3 x \, \bar{T} . \tag{3.187}$$

For a perfect fluid in a gravitational field, the stress-energy tensor reads

$$\bar{T} = \rho \vec{v} \otimes \vec{v} + P \, \mathbb{1}_3 + \bar{T}_{\mathrm{grav}} . \tag{3.188}$$

We integrate this over the complete volume of the fluid, use (3.178) and find

$$\frac{1}{2} \frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = \int \mathrm{d}^3 x \, \rho \vec{v} \otimes \vec{v} + \mathbb{1}_3 \int \mathrm{d}^3 x \, P + U . \tag{3.189}$$

This is the tensor virial theorem for a pefect fluid in its most general form. If the system is stable, the left-hand side vanishes, and a relation between the kinetic-energy tensor

$$K = \frac{1}{2} \int \mathrm{d}^3 x \, \rho \vec{v} \otimes \vec{v} , \tag{3.190}$$

the potential-energy tensor U^{ij} and the volume-integrated pressure remains,

$$2K = -\mathbb{1}_3 \int \mathrm{d}^3 x \, P - U . \tag{3.191}$$

3.3.6 Transformation to cylindrical or spherical coordinates

It is convenient in many applications of hydrodynamics to use coordinates other than Cartesian ones, in particular when systems with axial or spherical symmetry are to be studied. Then, of course, the spatial differential operators need to be transformed accordingly, but there is one more aspect of the transformation that needs to be taken into accout.

In cylindrical coordinates (r, φ, z), the basis vectors expressed in Cartesian coordinates are

$$\hat{e}_r = \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix} , \quad \hat{e}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} , \quad \hat{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} . \tag{3.204}$$

Since the position vector is $\vec{x} = r\hat{e}_r + z\hat{e}_z$, the velocity is

$$\vec{v} = \dot{r}\hat{e}_r + r\dot{\hat{e}}_r + \dot{z}\hat{e}_z = \dot{r}\hat{e}_r + r\dot{\varphi}\hat{e}_\varphi + \dot{z}\hat{e}_z , \tag{3.205}$$

Example: Virial theorem applied to a homogeneous sphere

To illustrate the power of the virial theorem to find out about the equilibrium state of a perfect fluid in a gravitational field, suppose we have a homogeneous sphere of density ρ, mass M and radius R which is macroscopically at rest, $v^i = 0$. The kinetic energy tensor vanishes, $K^{ij} = 0$. The fluid is assumed to have an ideal equation of state,

$$P = \frac{\rho}{m} k_\mathrm{B} T \ , \tag{3.192}$$

with a constant temperature throughout. Then,

$$\int \mathrm{d}^3 x\, P = \frac{M}{m} k_\mathrm{B} T \ . \tag{3.193}$$

By (3.179), the trace of the potential-energy tensor is

$$\begin{aligned}\mathrm{Tr}\, U &= \frac{1}{2} \int \mathrm{d}^3 x\, \rho\phi = -\frac{4\pi G\rho}{2} \int_0^R \frac{M(r)}{r} r^2 \mathrm{d}r = -\frac{3G}{10} \left(\frac{4\pi}{3}\right)^2 \rho^2 R^5 \\ &= -\frac{3}{10} \frac{GM^2}{R} \ .\end{aligned} \tag{3.194}$$

The trace of the tensor virial theorem (3.191) thus implies the relation

$$\frac{k_\mathrm{B} T}{m} = \frac{1}{10} \frac{GM}{R} \tag{3.195}$$

between the mass, the radius and the temperature of the sphere in equilibrium. Its so-called virial radius is

$$R = \frac{1}{10} \frac{GMm}{k_\mathrm{B} T} \ . \tag{3.196}$$

Suppose now that the sphere is rotating slowly like a solid body. The rotation needs to be slow to ensure that the body can still be assumed to be spherical. With a constant angular velocity $\vec\omega$, the velocity field is

$$\vec v = \vec\omega \times \vec r \ , \quad \vec v^{\,2} = \omega^2 r^2 \sin^2\theta \tag{3.197}$$

if we arrange the z axis of the coordinate system to be parallel to the angular velocity $\vec\omega$ and θ is the usual polar angle. The trace of the kinetic-energy tensor (3.190) is

$$\begin{aligned}\mathrm{Tr}\, K &= \frac{2\pi}{2} \rho\omega^2 \int_0^R r^4 \mathrm{d}r \int_0^\pi \sin^2\theta \sin\theta \mathrm{d}\theta \\ &= \frac{\pi}{5} \rho\omega^2 R^5 \int_{-1}^1 \left(1 - \mu^2\right) \mathrm{d}\mu \\ &= \frac{M}{5} \omega^2 R^2 \ .\end{aligned} \tag{3.198}$$

The trace of the tensor virial theorem (3.191) now gives the cubic equation

$$\frac{2}{5} \omega^2 R^3 + \frac{3}{m} k_\mathrm{B} T R - \frac{3GM}{10} = 0 \ . \tag{3.199}$$

◀

Example: Virial theorem applied to a slowly rotating, fln shomogeneous sphere

For a slowly rotating sphere, R will deviate little from the virial radius (3.196) of the sphere at rest, which we now call R_0 to write

$$R = R_0 + \delta R = R_0 \left(1 + \frac{\delta R}{R_0}\right) . \tag{3.200}$$

To lowest order in the small quantities ω^2 and δR, (3.200) can be approximated by

$$\frac{2}{5}\omega^2 R_0^3 + \frac{3}{m} k_B T R_0 \left(1 + \frac{\delta R}{R_0}\right) - \frac{3GM}{10} = 0 . \tag{3.201}$$

With R_0 from (3.196), we can further simplify this equation to

$$\delta R = -\frac{4}{3}\frac{\omega^2 R_0^4}{GM} . \tag{3.202}$$

Therefore, if the temperature of the fluid in the rotating sphere is the same as in the non-rotating sphere, its virial radius is reduced because the centrifugal force partly stabilises the body against gravity, allowing the body to be smaller. We can even set $T = 0$ in (3.199) and find

$$R = \left(\frac{3GM}{4\omega^2}\right)^{1/3} \tag{3.203}$$

for the radius of a stable, self-gravitating, rotating sphere. ◂

where $\dot{\hat{e}}_r = \dot{\varphi}\hat{e}_\varphi$ was inserted. We read off the velocity components

$$v_r = \dot{r} \;, \quad v_\varphi = r\dot{\varphi} \;, \quad v_z = \dot{z} \tag{3.206}$$

and write the acceleration as

$$\vec{a} = \dot{v}_r\hat{e}_r + \dot{v}_\varphi\hat{e}_\varphi + \dot{v}_z\hat{e}_z + v_r\dot{\hat{e}}_r + v_\varphi\dot{\hat{e}}_\varphi \;. \tag{3.207}$$

Since the time derivatives of the unit vectors $\hat{e}_r$ and $\hat{e}_\varphi$ are

$$\dot{\hat{e}}_r = \frac{v_\varphi}{r}\hat{e}_\varphi \;, \quad \dot{\hat{e}}_\varphi = -\frac{v_\varphi}{r}\hat{e}_r \;, \tag{3.208}$$

we can immediately identify the acceleration components

$$a_r = \dot{v}_r - \frac{v_\varphi^2}{r} \;, \quad a_\varphi = \dot{v}_\varphi + \frac{v_r v_\varphi}{r} \;, \quad a_z = \dot{v}_z \;. \tag{3.209}$$

Therefore, the components of the acceleration cannot simply be written as time derivatives of the velocity, but acquire additional terms. The expressions (3.209) imply that, in cylinder coordinates, the components in (r, φ, z) direction of the total time derivative on the left-hand side of Euler's equation needs to be augmented as

$$\mathrm{d}_t v_r \to \mathrm{d}_t v_r - \frac{v_\varphi^2}{r} \;, \quad \mathrm{d}_t v_\varphi \to \mathrm{d}_t v_\varphi + \frac{v_r v_\varphi}{r} \;, \quad \mathrm{d}_t v_z \to \mathrm{d}_t v_z \;. \tag{3.210}$$

?

Convince yourself by your own calculation of the expressions (3.208) and (3.213) for the time derivatives of the unit vectors.

In much the same way, we proceed for spherical polar coordinates (r, θ, φ), for which the basis vectors are

$$\hat{e}_r = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix} , \quad \hat{e}_\theta = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} , \quad \hat{e}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} . \tag{3.211}$$

Since $\dot{\hat{e}}_r = \dot{\theta}\hat{e}_\theta + \dot{\varphi}\sin\theta\hat{e}_\varphi$, the components of the velocity $\vec{v} = \dot{r}\hat{e}_r + r\dot{\hat{e}}_r$ are

$$v_r = \dot{r} \;, \quad v_\theta = r\dot{\theta} \;, \quad v_\varphi = r\dot{\varphi}\sin\theta \;. \tag{3.212}$$

We can thus write the time-derivatives of the unit vectors as

$$\begin{aligned} \dot{\hat{e}}_r &= \frac{v_\theta}{r}\hat{e}_\theta + \frac{v_\varphi}{r}\hat{e}_\varphi \;, \quad \dot{\hat{e}}_\theta = -\frac{v_\theta}{r}\hat{e}_r + \frac{v_\varphi}{r}\cot\theta\hat{e}_\varphi \;, \\ \dot{\hat{e}}_\varphi &= -\frac{v_\varphi}{r}\left(\hat{e}_r + \cot\theta\hat{e}_\theta\right) \end{aligned} \tag{3.213}$$

and immediately identify the components

$$\begin{aligned} a_r &= \dot{v}_r - \frac{v_\theta^2 + v_\varphi^2}{r} \;, \quad a_\theta = \dot{v}_\theta + \frac{v_r v_\theta}{r} - \frac{v_\varphi^2}{r}\cot\theta \;, \\ a_\varphi &= \dot{v}_\varphi + \frac{v_r v_\varphi}{r} + \frac{v_\theta v_\varphi}{r}\cot\theta \end{aligned} \tag{3.214}$$

of the acceleration.

In spherical coordinates, then, the left-hand side of Euler's equation needs to be transformed as

$$\begin{aligned}
\mathrm{d}_t v_r &\to \mathrm{d}_t v_r - \frac{v_\theta^2 + v_\varphi^2}{r} \\
\mathrm{d}_t v_\theta &\to \mathrm{d}_t v_\theta + \frac{v_r v_\theta}{r} - \frac{v_\varphi^2}{r} \cot\theta \\
\mathrm{d}_t v_\varphi &\to \mathrm{d}_t v_\varphi + \frac{v_r v_\varphi}{r} + \frac{v_\theta v_\varphi}{r} \cot\theta \,. \qquad (3.215)
\end{aligned}$$

The total time derivatives in the transformations (3.210) and (3.215) remain formally unchanged,

$$\mathrm{d}_t v_i = \left(\partial_t + \vec{v} \cdot \vec{\nabla}\right) v_i \,, \qquad (3.216)$$

but the gradient operator $\vec{\nabla}$ needs to be expressed in the respective coordinate basis.

Example: Hydrodynamic equations in cylinder coordinates

To give one specific example, we express the continuity and Euler equations for ideal hydrodynamics in cylinder coordinates (r, φ, z). Since the gradient and the divergence are

$$\vec{\nabla} f = \hat{e}_r \partial_r + \frac{\hat{e}_\varphi}{r} \partial_\varphi + \hat{e}_z \partial_z \quad \text{and} \quad \vec{\nabla} \cdot \vec{f} = \frac{1}{r} \partial_r (r f_r) + \frac{1}{r} \partial_\varphi f_\varphi + \partial_z f_z \,, \quad (3.217)$$

the continuity equation transforms to

$$\partial_t \rho + \frac{1}{r} \partial_r (r \rho v_r) + \frac{1}{r} \partial_\varphi (\rho v_\varphi) + \partial_z (\rho v_z) = 0 \,, \qquad (3.218)$$

while the components of Euler's equation turn into

$$\begin{aligned}
\partial_t v_r + \left(\vec{v} \cdot \vec{\nabla}\right) v_r - \frac{v_\varphi^2}{r} &= -\partial_r \left(\frac{P}{\rho} + \phi\right) , \\
\partial_t v_\varphi + \left(\vec{v} \cdot \vec{\nabla}\right) v_\varphi + \frac{v_r v_\varphi}{r} &= -\frac{1}{r} \partial_\varphi \left(\frac{P}{\rho} + \phi\right) , \\
\partial_t v_z + \left(\vec{v} \cdot \vec{\nabla}\right) v_z &= -\partial_z \left(\frac{P}{\rho} + \phi\right) , \qquad (3.219)
\end{aligned}$$

with the representation of $\vec{\nabla}$ to be taken from (3.217). ◂

Problems

1. Young stars often form in the centre of a thin accretion disk whose height is much smaller than its radius. If the mass of the central object M is much larger than the disk's mass, the gas particles move on approximately Keplerian orbits which are almost circular.

   (a) What is the velocity v of a gas particle as a function of the radius r? Determine also the divergence of the velocity field.

(b) Calculate the components of the velocity tensor

$$v_{ij} = \frac{1}{2}\left(\partial_j v_i - \partial_i v_j\right) \tag{3.220}$$

for the Keplerian disk.

3.4 Flows under specific circumstances

In this section, the hydrodynamical equations are applied to a variety of different flows. We begin with a perturbative analysis to derive the equation (3.226) for sound waves, identifying the expression (3.228) for the sound speed. Following the introduction of polytropic equations of state, we discuss hydrostatic equilibrium configurations and derive the Lane-Emden equation (3.259). Vorticity and circulation are defined next in the derivation of Kelvin's circulation theorem (3.280). Then, we demonstrate Bernoulli's law (3.286) for stationary flows by integration of Euler's equation and apply it to Bondi's problem of spherical accretion, leading to the relations (3.308) between velocity and radius in polytropic or isothermal flows. Next, we extend Bernoulli's law to non-stationary, but irrotational flows in (3.315). Viscous flows are briefly discussed at the end of the section. We begin with the diffusion of vorticity (3.317), define the Reynolds number (3.321) and conclude with viscous flow through a pipe, leading to the Hagen-Poiseulle law (3.329).

3.4.1 Sound waves

We begin with an ideal fluid for which we assume that a solution of the hydrodynamical equations is already given. This solution may consist of functions ρ_0, $\vec{v}_0$ and P_0, with the subscript 0 indicating that these functions are considered as a fixed, given, so-called background solution. We transform into the rest frame of this background solution and can thus assume $\vec{v}_0 = 0$. Then, we perturb this solution by small amounts $\delta\rho$, $\delta\vec{v}$ and δP, insert the perturbed solution

$$\rho = \rho_0 + \delta\rho \;, \quad \vec{v} = \delta\vec{v} \;, \quad P = P_0 + \delta P \tag{3.221}$$

into the continuity- and Euler equations and keep only terms up to first order in the perturbations. This procedure, which is typical for a perturbative analysis, results in

$$\partial_t(\rho_0 + \delta\rho) + \vec{\nabla}\cdot(\rho_0\delta\vec{v}) = 0 \;, \quad \partial_t\delta\vec{v} + \frac{\vec{\nabla}(P_0 + \delta P)}{\rho_0 + \delta\rho} = 0 \;. \tag{3.222}$$

Typically, the background solution is smooth on the length scale of the perturbations. If we may assume this, we can neglect gradients of ρ_0 and P_0 as well as $\partial_t\rho_0$ and continue writing

$$\partial_t\delta\rho + \rho_0\vec{\nabla}\cdot\delta\vec{v} = 0 \;, \quad \partial_t\delta\vec{v} + \frac{\vec{\nabla}\delta P}{\rho_0} = 0 \;. \tag{3.223}$$

We further relate the gradient of the pressure perturbation to the gradient of the density perturbation by

$$\vec{\nabla}\delta P = \frac{\partial P}{\partial \rho}\vec{\nabla}\delta\rho =: c_{\mathrm{s}}^2\vec{\nabla}\delta\rho \,, \tag{3.224}$$

introducing the abbreviation c_{s}^2 for the partial derivative of the pressure with respect to the density. Equations (3.223) then become

$$\partial_t\delta\rho + \rho_0\vec{\nabla}\cdot\delta\vec{v} = 0 \,, \quad \rho_0\partial_t\delta\vec{v} + c_{\mathrm{s}}^2\vec{\nabla}\delta\rho = 0 \,. \tag{3.225}$$

Taking a further time derivative of the first equation and the divergence of the second equation allows us to eliminate the velocity perturbation altogether and express the density perturbation as

$$\boxed{\partial_t^2\delta\rho - c_{\mathrm{s}}^2\vec{\nabla}^2\delta\rho = 0} \,. \tag{3.226}$$

This is a d'Alembert equation for the density contrast,

$$\Box\delta\rho = 0 \,, \tag{3.227}$$

in which c_{s} appears as the characteristic velocity. The solutions of (3.227) are linear density waves, accompanied by waves in the velocity perturbation. Such waves are sound waves, and

$$\boxed{c_{\mathrm{s}} = \left(\frac{\partial P}{\partial \rho}\right)^{1/2}} \tag{3.228}$$

is the sound speed. The derivative in (3.228) has to be taken at constant entropy.

> **?**
> Why would the sound speed (3.228) have to be determined at constant entropy? Is this necessarily so? What assumption may enter here?

The solutions of the d'Alembert equation can be expanded into plane, mono-"chromatic" waves. Let

$$\delta\rho = a\mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)} \,, \quad \delta\vec{v} = \vec{b}\mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)} \tag{3.229}$$

be such waves with wave vector $\vec{k}$ and frequency ω for the density and velocity perturbations. Inserting them into the d'Alembert equation gives the dispersion relation

$$k^2 = \frac{\omega^2}{c_{\mathrm{s}}^2} \tag{3.230}$$

familiar from electrodynamics, but with the sound speed in place of the light speed. The second equation (3.225), however, gives

$$\omega\rho_0\vec{b} = c_{\mathrm{s}}^2 a\vec{k} \,. \tag{3.231}$$

The amplitude $\vec{b}$ of the velocity perturbation is thus oriented with the wave vector $\vec{k}$, showing that $\delta\vec{v}$ is longitudinal.

3.4.2 Polytropic equation of state

We have noticed earlier that the equations of hydrodynamics are a set of five equations (one each for the conservation of the mass density, its internal energy and each of its momentum components) for six quantities, namely the density, the pressure, the internal energy or temperature of the fluid and its macroscopic velocity. One equation is missing. Typically, an equation of state is chosen for this purpose, that is an equation relating the pressure to the other fluid properties, such as the density and the temperature.

In astrophysics, it is frequently appropriate to assume the so-called polytropic relation between pressure and density,

$$P(\rho) = P_0 \left(\frac{\rho}{\rho_0}\right)^{\gamma} , \tag{3.232}$$

which can be derived for any fluid under adiabatic conditions. To see this, consider the first law of thermodynamics, $\delta Q = \mathrm{d}E + P\mathrm{d}V$. If no heat is exchanged, $\delta Q = 0$, and

$$\mathrm{d}E = c_\mathrm{v}\mathrm{d}T = -P\mathrm{d}V . \tag{3.233}$$

The enthalpy is obtained from the internal energy by the Legendre transform

$$H = E + PV , \quad \mathrm{d}H = \mathrm{d}E + P\mathrm{d}V + V\mathrm{d}P . \tag{3.234}$$

?

Why are infinitesimal changes of the internal energy and the enthalpy given by $\mathrm{d}E = c_\mathrm{v}\mathrm{d}T$ and $\mathrm{d}H = c_\mathrm{p}\mathrm{d}T$, respectively, with c_v and c_p being the heat capacities at constant volume or pressure?

Under adiabatic conditions, therefore,

$$\mathrm{d}H = c_\mathrm{p}\mathrm{d}T = V\mathrm{d}P . \tag{3.235}$$

If we now divide (3.235) by (3.233), the temperature differential $\mathrm{d}T$ cancels, and we find

$$\frac{c_\mathrm{p}}{c_\mathrm{v}} = \gamma = -\frac{V}{P}\frac{\mathrm{d}P}{\mathrm{d}V} , \tag{3.236}$$

where γ is defined to be the adiabatic index. Separating variables leads immediately to

$$\frac{\mathrm{d}P}{P} = -\gamma\frac{\mathrm{d}V}{V} , \tag{3.237}$$

or $P \propto V^{-\gamma} \propto \rho^{\gamma}$, which is already the polytropic relation (3.232). Notice in particular that we have nowhere used the assumption of an ideal gas. The entire derivation is based on the adiabatic condition that the fluid does not exchange heat with its environment. If we can additionally treat the fluid as an ideal gas, we have $PV \propto T$ and conclude

$$PV^{\gamma} = (PV)V^{\gamma-1} \propto TV^{\gamma-1} = \text{const.} \tag{3.238}$$

For an ideal gas, the polytropic relation (3.232) thus implies

$$T = T_0 \left(\frac{\rho}{\rho_0}\right)^{\gamma-1} . \tag{3.239}$$

The sound speed in a polytropic fluid is easily derived. We have to take the derivative of the pressure with respect to the density at constant entropy, but the polytropic relation has already been derived assuming that entropy is constant. It is therefore justified to write

$$c_\mathrm{s}^2 = \frac{\partial P}{\partial \rho} = \gamma\frac{P}{\rho} = c_\mathrm{s0}^2 \left(\frac{\rho}{\rho_0}\right)^{\gamma-1} . \tag{3.240}$$

For the enthalpy, we begin from (3.235) to derive the enthalpy per unit mass, $\tilde{h}$. Since the volume per unit mass is simply ρ^{-1}, we must have

$$\tilde{h} = \int \frac{\mathrm{d}P}{\rho} = \frac{\gamma}{\gamma-1}\frac{P}{\rho} = \frac{c_\mathrm{s}^2}{\gamma-1} \; . \tag{3.241}$$

The relations (3.240) and (3.241) are frequently used and often very convenient in discussions of astrophysical fluid flows.

Let us briefly remark on entropy here before continuing with the discussion of hydrodynamical flows under specific circumstances. The first law of thermodynamics states

$$T\mathrm{d}S = \mathrm{d}E + P\mathrm{d}V = c_\mathrm{v}\mathrm{d}T + P\mathrm{d}V = c_\mathrm{v}\mathrm{d}T + \mathrm{d}(PV) - V\mathrm{d}P \; , \tag{3.242}$$

where c_v is again the heat capacity at constant volume. Dividing by T, using the equation of state $PV = Nk_\mathrm{B}T$ for an ideal gas and the relation

$$c_\mathrm{p} - c_\mathrm{v} = Nk_\mathrm{B} \tag{3.243}$$

between the heat capacities c_p and c_v at constant pressure and constant volume, respectively, we transform (3.242) into

$$\mathrm{d}S = c_\mathrm{p}\frac{\mathrm{d}T}{T} - (c_\mathrm{p} - c_\mathrm{v})\frac{\mathrm{d}P}{P} \; . \tag{3.244}$$

Recalling the adiabatic index $\gamma = c_\mathrm{p}/c_\mathrm{v}$, we have

$$\mathrm{d}S = c_\mathrm{v}\left[\gamma\frac{\mathrm{d}T}{T} - (\gamma-1)\frac{\mathrm{d}P}{P}\right] , \tag{3.245}$$

from which we can infer the derivatives

$$\left(\frac{\partial S}{\partial T}\right)_P = \gamma\frac{c_\mathrm{v}}{T} \quad \text{and} \quad \left(\frac{\partial S}{\partial P}\right)_T = -(\gamma-1)\frac{c_\mathrm{v}}{P} \tag{3.246}$$

for the entropy with respect to T at constant P, and vice versa. We shall need these relations in the derivation of the convective instability below.

Caution Recall the Maxwell relation

$$\left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P$$

which, when evaluated for an ideal gas, results in

$$\left(\frac{\partial S}{\partial P}\right)_T = -\frac{Nk_\mathrm{B}}{P} = -\frac{c_\mathrm{p} - c_\mathrm{v}}{P} \; .$$

◀

From the ideal gas equation written in the form

$$T = \frac{PV}{Nk_\mathrm{B}} = \frac{P}{\rho}\frac{\bar{m}}{k_\mathrm{B}} \tag{3.247}$$

with the mean particle mass $\bar{m}$, we immediately infer that

$$\frac{\mathrm{d}T}{T} = \frac{\mathrm{d}P}{P} - \frac{\mathrm{d}\rho}{\rho} \; , \tag{3.248}$$

and insert this expression into (3.244) to find

$$\mathrm{d}S = c_\mathrm{v}\frac{\mathrm{d}P}{P} - c_\mathrm{p}\frac{\mathrm{d}\rho}{\rho} \; . \tag{3.249}$$

The derivatives of the entropy with respect to P at constant ρ and vice versa are thus

$$\left(\frac{\partial S}{\partial P}\right)_\rho = \frac{c_\mathrm{v}}{P} \quad \text{and} \quad \left(\frac{\partial S}{\partial \rho}\right)_P = -\frac{c_\mathrm{p}}{\rho} \; . \tag{3.250}$$

We shall return to these relations in the discussion of the thermal instability.

Finally, it is instructive to conclude from (3.249) that the entropy as a function of pressure and density is

$$S(P,\rho) = c_\mathrm{v} \ln\left[\frac{P}{P_0}\left(\frac{\rho_0}{\rho}\right)^\gamma\right] . \tag{3.251}$$

For a polytropic gas with $P \propto \rho^\gamma$, the entropy is manifestly constant, as it should be by construction.

3.4.3 Hydrostatic equilibrium

We begin our study of hydrodynamical flows under specific, generally simplifying conditions with a fluid in hydrostatic equilibrium. In a static situation, the flow velocity vanishes, $\vec{v} = 0$, and the Navier-Stokes equation (3.148) shrinks to

$$\vec{\nabla} P = -\rho\vec{\nabla}\Phi . \tag{3.252}$$

Taking the curl of this equation, we immediately see that

$$\vec{\nabla}\rho \times \vec{\nabla}\Phi = 0 \tag{3.253}$$

because the curl of a gradient vanishes identically. The gradients of the gravitational potential and of the density must therefore be parallel to each other, which means that the equipotential surfaces, i.e. the surfaces of constant potential, must also be the surfaces of constant density. In hydrostatic equilibrium, the shape of the fluid body thus adapts to the shape of the gravitational potential.

Taking the divergence of the hydrostatic equation, we can use Poisson's equation to write

$$\vec{\nabla} \cdot \left(\frac{\vec{\nabla} P}{\rho}\right) = -4\pi G\rho . \tag{3.254}$$

Once an equation of state is chosen for the fluid, i.e. a relation between the pressure P and the density ρ, this equation determines the configuration of the fluid density in its own gravitational field. Let us suppose that the pressure satisfies the polytropic relation, and restrict the discussion to spherically-symmetric configurations. Then,

$$\frac{1}{r^2}\partial_r\left(r^2\frac{\partial_r P}{\rho}\right) = \frac{c_{\mathrm{s}0}^2}{r^2}\partial_r\left[r^2\left(\frac{\rho}{\rho_0}\right)^{\gamma-1}\partial_r\rho\right] = -4\pi G\rho . \tag{3.255}$$

Instead of the adiabatic index, we now introduce the polytropic index n by defining

$$\gamma - 1 = \frac{1}{n} . \tag{3.256}$$

Moreover, we introduce a function θ to describe the density as

$$\frac{\rho}{\rho_0} = \theta^n , \tag{3.257}$$

define a characteristic radius

$$r_0 = \left(\frac{n c_{\mathrm{s}0}^2}{4\pi G\rho_0}\right)^{1/2} \tag{3.258}$$

and use that to introduce the dimension-less radial coordinate $x = r/r_0$. These operations leave (3.255) in the dimension-less form

$$\frac{1}{x^2}\partial_x\left(x^2\partial_x\theta\right) = -\theta^n \ , \tag{3.259}$$

which is called the Lane-Emden equation.

?

Independently carry out all steps leading from the hydrostatic equation (3.252) to the Lane-Emden equation (3.259).

Given a polytropic index n, it can be solved with the boundary conditions $\partial_x\theta = 0$ and $\theta = 1$ at $x = 0$ to return the density profile of a polytropic, self-gravitating gas sphere. Expanding the differential operator in (3.259), the Lane-Emden equation reads

$$\theta'' + \frac{2}{x}\theta' + \theta^n = 0 \ . \tag{3.260}$$

Example: Solutions of the Lane-Emden equation

Analytic solutions for the Lane-Emden equation exist for $n = 0$, $n = 1$ and $n = 5$. For $n = 0$, direct integration of (3.259) results in

$$\theta = -\frac{x^2}{6} - \frac{A}{x} + B \tag{3.261}$$

with two integration constants A and B. The boundary conditions require $A = 0$ for the solution to remain regular at the centre and $B = 1$ for θ to reach unity there. Thus,

$$\theta(x) = 1 - \frac{x^2}{6} \tag{3.262}$$

for $n = 0$. For $n = 1$, (3.260) is a spherical Bessel differential equation of order zero, which is solved by spherical Bessel function

$$\theta(x) = j_0(x) = \frac{\sin x}{x} \ . \tag{3.263}$$

Numerical solutions for the Lane-Emden equation with adiabatic indices $\gamma = 5/3$ (polytropic index $n = 3/2$) or $\gamma = 4/3$ ($n = 3$) are often used to model the internal structure of white dwarfs or other stars (Figure 3.4). ◂

Another interesting and illustrative example for systems in hydrostatic equilibrium is the case of a gas filled into a spherical gravitational potential well caused by the dominant dark matter. If the gas mass is overall negligible, the gravitational potential is given independently, and the gas just responds to it. This requires us to separate the gas density $\rho_{\rm gas}$ from the dark-matter density $\rho_{\rm DM}$ in the hydrostatic equation, thus

$$\frac{1}{r^2}\partial_r\left(\frac{r^2}{\rho_{\rm gas}}\partial_r P\right) = -4\pi G\rho_{\rm DM} \ . \tag{3.264}$$

With the equation of state for an ideal gas,

$$P = \frac{\rho_{\rm gas}}{m}k_{\rm B}T \ , \tag{3.265}$$

where m is the (mean) mass of a gas particle, we find by integrating once

$$\frac{r^2}{\rho_{\rm gas}}\frac{k_{\rm B}}{m}\partial_r\left(\rho_{\rm gas}kT\right) = -4\pi G\int_0^r r'^2{\rm d}r'\,\rho_{\rm DM} = -GM_{\rm DM}(r) \ , \tag{3.266}$$

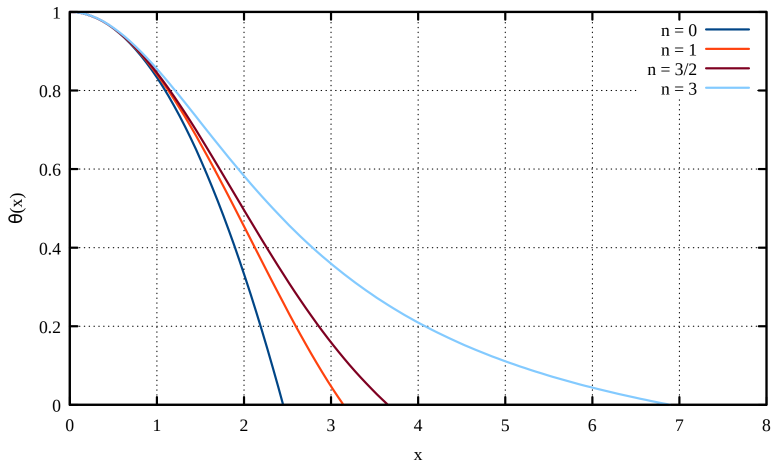

Figure 3.4 Solutions of the Lane-Emden equation are shown for different choices of the polytropic index. The curves are displayed up to their first root only.

where $M_{\rm DM}(r)$ is the dark-matter mass enclosed by a sphere of radius r. Solving this equation for the dark-matter mass shows how it is related to the logarithmic gradients of temperature and gas density,

$$M_{\rm DM}(r) = -\frac{rk_{\rm B}T}{mG}\left(\frac{{\rm d}\ln\rho_{\rm gas}}{{\rm d}\ln r} + \frac{{\rm d}\ln T}{{\rm d}\ln r}\right) . \tag{3.267}$$

This equation is often applied to find mass estimates for galaxy clusters. There, the two logarithmic gradients can be inferred from X-ray observations of the hot intracluster gas.

3.4.4 Vorticity and Kelvin's circulation theorem

We shall now give up the hydrostatic assumption, but still neglect any dissipative effects, such as viscous friction and heat conduction. In the Navier-Stokes equation (3.148), we therefore set $\eta = 0 = \zeta$, and $\kappa = 0$ in the energy-conservation equation. We then also know from (3.163) that entropy is conserved under such circumstances. Momentum conservation is then expressed by Euler's equation

$$\partial_t\vec{v} + \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v} + \frac{\vec{\nabla}P}{\rho} + \vec{\nabla}\Phi = 0 . \tag{3.268}$$

The identity

$$\vec{v}\times\left(\vec{\nabla}\times\vec{v}\right) = \vec{\nabla}\left(\frac{v^2}{2}\right) - \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v} \tag{3.269}$$

enables us to replace the convective velocity derivative $(\vec{v}\cdot\vec{\nabla})\vec{v}$ in (3.268) to obtain

$$\partial_t\vec{v} - \vec{v}\times\left(\vec{\nabla}\times\vec{v}\right) = -\vec{\nabla}\left(\frac{v^2}{2}\right) - \frac{\vec{\nabla}P}{\rho} - \vec{\nabla}\Phi . \tag{3.270}$$

The curl of the velocity,

$$\vec{\Omega} := \vec{\nabla}\times\vec{v} , \tag{3.271}$$

is called the *vorticity* of the flow. If we take the curl of Euler's equation in its form (3.270), we find the evolution equation for the vorticity

$$\partial_t\vec{\Omega} = \vec{\nabla}\times\left(\vec{v}\times\vec{\Omega}\right) + \frac{\vec{\nabla}\rho\times\vec{\nabla}P}{\rho^2} \tag{3.272}$$

since the curl of the gradients vanishes identically. If the pressure P is a function of ρ only, as for example in a polytropic fluid, the gradients of P and ρ must align because then

$$\vec{\nabla}P = \frac{\mathrm{d}P}{\mathrm{d}\rho}\vec{\nabla}\rho \quad\Rightarrow\quad \vec{\nabla}P\times\vec{\nabla}\rho = 0\ . \tag{3.273}$$

For such *barotropic* fluids, the vorticity equation simplifies to

$$\frac{\partial\Omega}{\partial t} = \vec{\nabla}\times\left(\vec{v}\times\vec{\Omega}\right)\ . \tag{3.274}$$

?

What does $\partial_t\vec{\Omega} = 0$ imply for the solution(s) of the vorticity equation (3.274) for barotropic fluids?

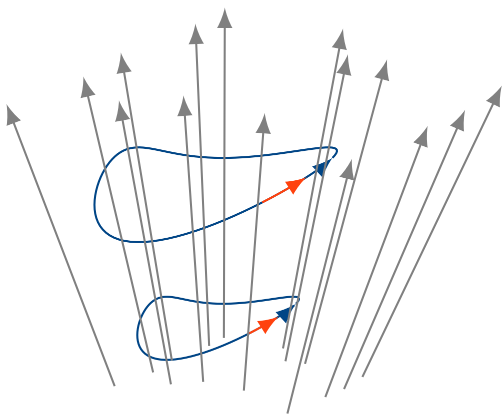

Figure 3.5 Illustration of Kelvin's circulation theorem: The circulation of the velocity field in an inviscid fluid is conserved.

Having derived an evolution equation for the vorticity, we now consider the so-called *circulation*, which is the line integral over the velocity along closed curves swimming with the fluid flow,

$$\Gamma := \oint_C \vec{v}\cdot\mathrm{d}\vec{l}\,. \tag{3.275}$$

We are interested in the total change with time of the circulation embedded into the flow (Figure 3.5). We must therefore take into consideration that the contour C is deformed by the flow. The total time derivative of Γ consists of the change of the velocity field within the contour, plus the change of the contour itself. For a more transparent notation, we write the infinitesimal path length $\mathrm{d}\vec{l}$ as a difference $\delta\vec{r}$ of the position vectors pointing at the beginning and the end of $\mathrm{d}\vec{l}$. Accordingly, we write

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = \oint_C \frac{\mathrm{d}\vec{v}}{\mathrm{d}t}\cdot\mathrm{d}\vec{l} + \oint_C \vec{v}\cdot\frac{\mathrm{d}\delta\vec{r}}{\mathrm{d}t}\ . \tag{3.276}$$

In the first term on the right-hand side, we expand the total time derivative of the velocity into

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \partial_t\vec{v} + \left(\vec{v}\cdot\vec{\nabla}\right)\vec{v} = \partial_t\vec{v} + \vec{\nabla}\left(\frac{v^2}{2}\right) - \vec{v}\times\Omega\ . \tag{3.277}$$

The line integral suggests taking the curl and applying Stokes' law. The curl of (3.277) is

$$\frac{\mathrm{d}\vec{\Omega}}{\mathrm{d}t} = \partial_t\vec{\Omega} - \vec{\nabla}\times\left(\vec{v}\times\vec{\Omega}\right) = 0 \tag{3.278}$$

according to the vorticity equation (3.274), hence the first term of the total time derivative (3.276) of the circulation vanishes. The second term is

$$\oint_C \vec{v}\cdot\frac{\mathrm{d}\delta\vec{r}}{\mathrm{d}t} = \oint_C \vec{v}\cdot\delta\vec{v} = \oint_C \vec{\nabla}\left(\frac{v^2}{2}\right)\cdot\mathrm{d}\vec{l} = 0\ , \tag{3.279}$$

which also vanishes because an integral along a closed loop of a gradient field must vanish. The circulation is thus conserved in a barotropic, ideal fluid,

$$\boxed{\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = 0}\ , \tag{3.280}$$

which is Kelvin's circulation theorem.

3.4.5 Bernoulli's constant

If the fluid is not static, but the flow is stationary, all partial derivatives with respect to time will vanish. In such cases, flow lines can be introduced as the integral curves of the velocity field. Quite obviously, the flow lines must obey the equations

$$\frac{\mathrm{d}x}{v_x} = \mathrm{d}t = \frac{\mathrm{d}y}{v_y} = \frac{\mathrm{d}z}{v_z}\ . \tag{3.281}$$

In ideal fluids, we have seen that the specific entropy $\tilde{s}$ is constant because energy dissipation and heat flows do not occur. For a stationary flow, $\partial_t\tilde{s} = 0$ and

$$\frac{\mathrm{d}\tilde{s}}{\mathrm{d}t} = \partial_t\tilde{s} + \left(\vec{v}\cdot\vec{\nabla}\right)\tilde{s} = \left(\vec{v}\cdot\vec{\nabla}\right)\tilde{s} = 0\ , \quad \text{thus} \quad \left(\vec{v}\cdot\vec{\nabla}\right)\tilde{s} = 0 \tag{3.282}$$

The specific entropy must therefore be constant along flow lines. Moreover, we have seen in (3.241) that the specific enthalpy per unit mass satisfies

$$\mathrm{d}\tilde{h} = \frac{\mathrm{d}P}{\rho} \tag{3.283}$$

under adiabatic conditions.

For stationary flows, $\partial_t\vec{v} = 0$, and Euler's equation in its form (3.270) implies

$$\frac{1}{2}\vec{\nabla}\left(\vec{v}^2\right) - \vec{v}\times\vec{\Omega} = -\frac{\vec{\nabla}P}{\rho} - \vec{\nabla}\Phi\ . \tag{3.284}$$

Let us now multiply (3.284) with the fluid velocity $\vec{v}$ to obtain the change of its terms with time along flow lines. The term containing the vector product

$\vec{v} \times \vec{\Omega}$ then vanishes because it is perpendicular to $\vec{v}$. The remaining terms can be combined under the gradient,

$$\vec{v} \cdot \vec{\nabla} \left(\frac{v^2}{2} + \tilde{h} + \Phi \right) = 0 \; . \tag{3.285}$$

This reveals that the term in parentheses,

$$\frac{v^2}{2} + \tilde{h} + \Phi =: B = \text{constant along flow lines} \tag{3.286}$$

must be a constant B along flow lines, which is called Bernoulli's constant. We have thus proven Bernoulli's very important and intuitive law for ideal flows. It shows that the specific kinetic energy of the flow, $v^2/2$, is not only balanced by the specific potential energy in the gravitational field, but also by the specific enthalpy. For example, if a gas flow is expanding as it propagates into a surrounding medium and against a gravitational field, the enthalpy term takes into account that the gas will have to exert pressure-volume work against the surrounding medium and thereby cool.

Example: The faucet

Bernoulli's law, together with the equation of continuity, are very powerful tools to study stationary fluid flows. Let us begin with water flowing from a faucet, accelerated by gravity (Figure 3.6). Everyday experience tells us that the diameter of the water shrinks as it falls. How exactly does the diameter depend on the height, and why?

Bernoulli's law tells us that the quantity

$$\frac{v^2}{2} + \tilde{h} + \Phi = \text{const} \tag{3.287}$$

along the flow lines of the water. Here, we can replace the gravitational potential by $\Phi = gz$ if z points vertically upwards, where g is the local gravitational acceleration. The pressure is set by the atmospheric pressure surrounding the water, the density can be assumed to be constant. Bernoulli's law then tells us that the water accelerates as it falls according to $v^2 = v_0^2 + 2g(h - z)$ if it is initially at rest at the height h.

To evaluate the continuity equation for a stationary flow, $\partial_t \rho = 0$, we integrate the divergence $\vec{\nabla} \cdot (\rho \vec{v}) = 0$ over an infinitesimally thin cylinder with cross section A whose axis is aligned with the water. Gauss' law then implies that

$$\rho v A = \text{const} = \rho v_0 A_0 \; , \tag{3.288}$$

from which we conclude that

$$A = \frac{A_0 v_0}{\sqrt{v_0^2 + 2g(h - z)}} \; . \tag{3.289}$$

The cross section of the water decreases as it falls from $z = h$ to $z = 0$. ◂

Example: The Laval nozzle

Completely analogous is the discussion of gas flowing through a nozzle whose cross section first decreases, then increases along the gas flow. For definiteness, the x axis of the coordinate system may point into the direction of the gas flow, and the cross section $A(x)$ is given. Continuity now demands

$$\rho v A(x) = \rho_0 v_0 A_0 \ , \tag{3.290}$$

while Bernoulli's law requires

$$\frac{v^2}{2} + \frac{c_s^2 - c_{s0}^2}{\gamma - 1} = \frac{v_0^2}{2} \ . \tag{3.291}$$

Dividing by the squared initial sound speed c_{s0}^2 gives the dimension-less equation

$$\frac{u^2}{2} + \frac{\alpha^{\gamma-1} - 1}{\gamma - 1} = \frac{u_0^2}{2} \ , \tag{3.292}$$

where the dimension-less density $\alpha := \rho/\rho_0$ was introduced. A similar operation brings the continuity equation into the form

$$\alpha u A = u_0 A_0 \ . \tag{3.293}$$

Let us now take the total differentials of both equations (3.292) and (3.293). This leads us to

$$u\mathrm{d}u + \alpha^{\gamma-2}\mathrm{d}\alpha = 0 \ , \quad \frac{\mathrm{d}\alpha}{\alpha} + \frac{\mathrm{d}u}{u} + \frac{\mathrm{d}A}{A} = 0 \ . \tag{3.294}$$

Eliminating $\mathrm{d}\alpha$ between these two equations leaves us with the equation

$$u\mathrm{d}u\left(1 - \frac{\alpha^{\gamma-1}}{u^2}\right) = u\mathrm{d}u\left(1 - \frac{1}{\mathcal{M}^2}\right) = \alpha^{\gamma-1}\frac{\mathrm{d}A}{A} \ , \tag{3.295}$$

where we have identified the squared local Mach number $\mathcal{M}^2 = u^2/\alpha^{\gamma-1}$. As long as the flow remains subsonic, $\mathcal{M} < 1$, the left-hand side is negative. The flow will continue to accelerate, $u\mathrm{d}u > 0$, if the cross section of the nozzle decreases, $\mathrm{d}A < 0$. This agrees with everyday experience: A gas flow through a narrowing pipe accelerates. However, the sign changes once the flow becomes supersonic, $\mathcal{M} > 1$. Then, for $u\mathrm{d}u$ to remain positive, the cross section of the nozzle must *increase*, $\mathrm{d}A > 0$! Otherwise, once the sonic point is reached, the gas will decelerate in narrowing nozzle. If the situation is arranged such that the sound speed is reached at the narrowest point of the nozzle, the flow will continue accelerating. This is the principle of the Laval nozzle, which is used for example in rocket engines (Figure 3.7). ◂

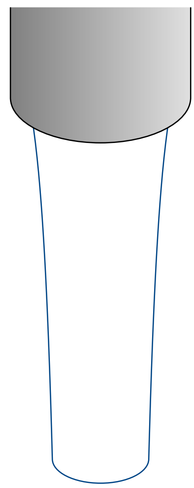

Figure 3.6 Water running from a faucet has a cross section determined by Bernoulli's law.

3.4.6 Bondi accretion

Completely analogous to the discussion of the faucet and the Laval nozzle is Bondi's accretion problem. The situation is as follows: A star or another point-like gravitating body of mass M is placed into a formerly homogeneous, extended gas cloud of density ρ_0 and pressure P_0. Driven by gravity, the gas will flow towards the star. How does it flow, and how much gas per unit time will the star accrete? Again, Bernoulli's law and the continuity equation provide the complete answer.

For a stationary, spherically-symmetric flow, the continuity equation reads

$$\frac{1}{r^2}\partial_r\left(r^2\rho v\right) = 0 \quad \Rightarrow \quad r^2\rho v = \text{const.} \tag{3.296}$$

The constant has the dimension g s^{-1} and therefore corresponds to the accretion rate, i.e. the rate at which matter flows onto the star. If we multiply (3.296) with 4π, we obtain the mass per unit time $\dot{M}$ flowing through the complete spherical surface,

$$4\pi r^2\rho v = -\dot{M} \; , \tag{3.297}$$

where the minus sign is introduced to express that the mass is flowing towards the star.

Bernoulli's law reads

$$\frac{v^2}{2} + \frac{c_s^2 - c_{s0}^2}{\gamma - 1} - \frac{GM}{r} = 0 \tag{3.298}$$

because the gas is assumed to be at rest far away from the star. This equation holds for adiabatic gas which can be treated as a polytrope. If the gas is isothermal and ideal rather than polytropic, its enthalpy per unit mass is

$$\tilde{h} = \int\frac{\mathrm{d}P}{\rho} = \frac{k_\mathrm{B}T}{m}\int\frac{\mathrm{d}\rho}{\rho} = c_{s0}^2\ln\left(\frac{\rho}{\rho_0}\right) \; , \tag{3.299}$$

Figure 3.7 An example for a de Laval nozzle is the Vulcain-II engine of an Ariane 5 rocket (Wikipedia, Creative Commons License)

and Bernoulli's law becomes

$$\frac{v^2}{2} + c_{s0}^2 \ln\left(\frac{\rho}{\rho_0}\right) - \frac{GM}{r} = 0 \qquad (3.300)$$

instead. We now divide both versions of Bernoulli's law by the unperturbed, squared sound speed c_{s0}^2, introduce the dimension-less velocity $u = v/c_{s0}$, the density $\alpha = \rho/\rho_0$, the so-called Bondi-radius

$$r_B = \frac{GM}{c_{s0}^2} \qquad (3.301)$$

and the dimension-less radius $x := r/r_B$. These substitutions leave our two versions of Bernoulli's equations in the convenient, dimension-less forms

$$\frac{u^2}{2} + \frac{\alpha^{\gamma-1} - 1}{\gamma - 1} - \frac{1}{x} = 0\,, \quad \frac{u^2}{2} + \ln\alpha - \frac{1}{x} = 0\,. \qquad (3.302)$$

The same substitutions turn the continuity equation into

$$x^2 \alpha u = \mu\,, \quad \mu := -\frac{\dot{M}}{4\pi r_B^2 \rho_0 c_{s0}}\,. \qquad (3.303)$$

The parameter μ is the accretion rate in units of the so-called Bondi accretion rate,

$$\dot{M}_B = 4\pi r_B^2 \rho_0 c_{s0}\,. \qquad (3.304)$$

We now have two equations, the continuity equation (3.303) and Bernoulli's law (3.302), for the two functions α and u. Eliminating α between them leaves one equation for the velocity u,

$$\frac{u^2}{2} + \frac{\left(\frac{\mu}{x^2 u}\right)^{\gamma-1} - 1}{\gamma - 1} - \frac{1}{x} = 0\,, \quad \frac{u^2}{2} + \ln\left(\frac{\mu}{x^2 u}\right) - \frac{1}{x} = 0\,. \tag{3.305}$$

But what accretion rates are possible? Does the flow turn supersonic somewhere? And if so, what happens? In order to see this, let us take complete differentials of the continuity equation,

$$\frac{2\mathrm{d}x}{x} + \frac{\mathrm{d}\alpha}{\alpha} + \frac{\mathrm{d}u}{u} = 0\,, \tag{3.306}$$

and of Bernoulli's law,

$$u\mathrm{d}u + \alpha^{\gamma-1}\frac{\mathrm{d}\alpha}{\alpha} + \frac{\mathrm{d}x}{x^2} = 0\,, \quad u\mathrm{d}u + \frac{\mathrm{d}\alpha}{\alpha} + \frac{\mathrm{d}x}{x^2} = 0\,, \tag{3.307}$$

and eliminate $\mathrm{d}\alpha/\alpha$ between them. For the polytropic gas, the squared sound speed is $c_\mathrm{s}^2 = c_{\mathrm{s}0}^2\alpha^{\gamma-1}$ according to (3.240), while it is constant $c_\mathrm{s}^2 = c_{\mathrm{s}0}^2$ for the isothermal gas. This leads to

$$u\mathrm{d}u\left(1 - \frac{1}{\mathcal{M}^2}\right) = \begin{cases} \dfrac{\mathrm{d}x}{x}\left(2\alpha^{\gamma-1} - \dfrac{1}{x}\right) & \text{polytropic} \\[2ex] \dfrac{\mathrm{d}x}{x}\left(2 - \dfrac{1}{x}\right) & \text{isothermal} \end{cases}\,, \tag{3.308}$$

where we have once more identified the squared Mach number $\mathcal{M}$ as before in (3.295) for the Laval nozzle.

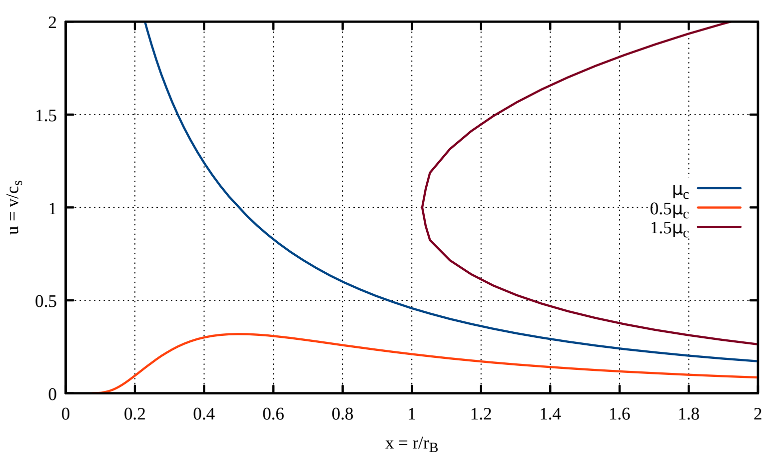

Figure 3.8 The radial velocity is shown as a function of radius for isothermal Bondi accretion. Velocity curves are given for three different accretion rates: the critical accretion rate μ_c in units of the Bondi accretion rate as well as 50 % more or less. The curve for $\mu_\mathrm{c} > 1$ is mathematically possible, but physically excluded because it corresponds to two velocities at the same radius.

This equation shows that there exists a critical radius, $x_\mathrm{c} = 1/2$ in the isothermal and $x_\mathrm{c} = \alpha^{1-\gamma}/2$ in the polytropic case, where the right-hand side vanishes. The left-hand side must then also vanish, which is either possible if the flow comes to a halt there, $u = 0$, if the velocity reaches a maximum, $\mathrm{d}u = 0$, or if the flow turns supersonic, $\mathcal{M} = 1$. What exactly happens, depends on the accretion rate (Figure 3.8). If the flow turns supersonic at the critical radius, $u^2 = \alpha^{\gamma-1}$ in the polytropic and $u = 1$ in the isothermal case, we can solve Bernoulli's equation (3.302) for the dimension-less density α there, obtaining

$$\alpha = \left(\frac{2}{5-3\gamma}\right)^{1/(\gamma-1)} \quad \text{(polytropic)}\,, \quad \alpha = \mathrm{e}^{3/2} \quad \text{(isothermal)}\,. \tag{3.309}$$

The continuity equation finally gives the critical accretion rate,

$$\mu_\mathrm{c} = \frac{1}{4}\left(\frac{2}{5-3\gamma}\right)^{(5-3\gamma)/(2(\gamma-1))}\,, \quad \mu_\mathrm{c} = \frac{\mathrm{e}^{3/2}}{4}\,. \tag{3.310}$$

For accretion rates smaller than μ_c, the flow speed reaches a subsonic maximum at the critical radius, corresponding to a gentle accretion flow that is everywhere subsonic. For accretion rates higher than μ_c, the solution is mathematically possible, but not physically: as Fig. 3.8 shows, the velocity then becomes double-valued where it exists, while two velocities at the same radius cannot exist in a fluid.

?

What would happen to the preceding calculation if the accretion rate $-\dot{M}$ from (3.297) would be set negative? What physical situation would this corresponed to?

3.4.7 Bernoulli's law for irrotational, non-stationary flows

We have derived Bernoulli's law for stationary flows before. It can be generalised to some degree for irrotational flows. For those, $\vec{\nabla} \times \vec{v} = \vec{\Omega} = 0$, which allows us to introduce a velocity potential ψ such that $\vec{v} = \vec{\nabla}\psi$. Euler's equation can then be written in the form

$$\partial_t \vec{\nabla}\psi + \vec{\nabla}\left(\frac{v^2}{2}\right) + \frac{\vec{\nabla}P}{\rho} + \vec{\nabla}\Phi = 0\,. \tag{3.311}$$

For adiabatic flows,

$$\vec{\nabla}\tilde{h} = \frac{\vec{\nabla}P}{\rho}\,, \tag{3.312}$$

hence we can infer from (3.311) that the function

$$\partial_t\psi + \frac{v^2}{2} + \tilde{h} + \Phi = B(t) \tag{3.313}$$

must be a function of time only. Since the velocity is given by a spatial gradient of ψ, we can gauge the velocity potential such that the right-hand side of (3.313) vanishes,

$$\psi \to \psi + \int \mathrm{d}t\, B(t) \tag{3.314}$$

and simplify (3.313) to Bernoulli's law for non-stationary, but irrotational flows,

$$\partial_t\psi + \frac{v^2}{2} + \tilde{h} + \Phi = 0\,. \tag{3.315}$$

3.4.8 Diffusion of vorticity

Let us now turn to simple examples of viscous flows. We begin with the Navier-Stokes equation (3.148) in the form

$$\partial_t \vec{v} + \vec{\nabla}\left(\frac{v^2}{2}\right) - \vec{v}\times\vec{\Omega} = \frac{1}{\rho}\left[-\vec{\nabla}P + \eta\vec{\nabla}^2\vec{v} + \left(\zeta + \frac{\eta}{3}\right)\vec{\nabla}\left(\vec{\nabla}\cdot\vec{v}\right)\right] \tag{3.316}$$

and take its curl. For simplicity, we assume that the flow is incompressible, $\vec{\nabla}\rho = 0$, and find

$$\partial_t\vec{\Omega} - \vec{\nabla}\times\left(\vec{v}\times\vec{\Omega}\right) = \nu\vec{\nabla}^2\vec{\Omega} \;, \tag{3.317}$$

where the kinematic viscosity

$$\nu := \frac{\eta}{\rho} \tag{3.318}$$

was introduced. Equation (3.317) relates a first-order partial time derivative to a second-order spatial derivative and is thus a diffusion equation for the vorticity. It shows how vorticity diffuses away in presence of viscosity.

3.4.9 The Reynolds Number

The kinematic viscosity has the dimension

$$\frac{\mathrm{g}}{\mathrm{cm\,s}}\,\frac{\mathrm{cm}^3}{\mathrm{g}} = \frac{\mathrm{cm}^2}{\mathrm{s}} \;, \tag{3.319}$$

that is, it is squared length over time. Suppose we scale all lengths with a typical length scale L, all velocities with a typical velocity V and all times with a time scale L/V in the vorticity equation (3.317). The expressions occuring would then scale as

$$\partial_t \to \frac{L}{V}\partial_t \;, \quad \partial_x \to L\partial_x \;, \quad \vec{v} \to \frac{\vec{v}}{V} \;, \quad \vec{\Omega} \to \vec{\Omega}\,\frac{L}{V} \;, \quad \nu \to \frac{\nu}{LV} \;, \tag{3.320}$$

such that all terms in (3.317) would be scaled by L^2/V^2 and thus become dimension-less. Therefore, if we characterise the viscosity by the dimension-less number

$$\frac{\nu}{LV} =: \frac{1}{\mathcal{R}} \;, \tag{3.321}$$

nothing in (3.317) reminds of the dimensions and the velocity of the flow. This shows that flows with the same Reynolds number $\mathcal{R}$ are scale-free. If lengths and velocities in a flow are stretched by factors L and V, respectively, the flow remains the same if the viscosity is simultaneously scaled by LV. The Reynolds number thus classifies such self-similar solutions of the flow equations. The transition to ideal fluids is characterised by $\mathcal{R} \to \infty$.

3.4.10 Hagen-Poiseulle flow

As one instructive example for a viscous flow, let us consider a viscous fluid running under the influence of a pressure gradient through a long, straight pipe.

Long means that its length L is much larger than its Radius R. We turn the coordinate system such that the symmetry axis of the pipe coincides with the x axis. The velocity $\vec{v}$ will then only have an x component which will itself only depend on the y and z coordinates perpendicular to the pipe. Since there are no other components of $\vec{v}$, we write $\vec{v} = v(y, z)\hat{e}_x$.

For a stationary flow, the continuity equation requires

$$\partial_x(\rho v) = v\partial_x\rho + \rho\partial_x v = 0 , \tag{3.322}$$

from where we read off that the density ρ will not depend on x either.

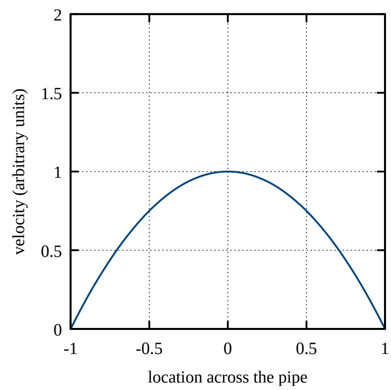

Figure 3.9 Illustration of the parabolic velocity profile across a pipe for Hagen-Poiseulle flow.

The Navier-Stokes equation shrinks to

$$\partial_x P = \eta\left(\partial_y^2 + \partial_z^2\right)v , \quad \partial_y P = 0 = \partial_z P \tag{3.323}$$

because the partial time derivative $\partial_t\vec{v}$ vanishes for a stationary flow, $(\vec{v}\cdot\vec{\nabla})v = v\partial_x v = 0$ since v does not depend on x, and $\vec{\nabla}\cdot\vec{v} = 0$ for the same reason. The second equation tells us that P is constant on planes perpendicular to the pipe. Since the right-hand side of the first equation cannot depend on x, neither can $\partial_x P$, hence $\partial_x P$ is constant along the pipe,

$$\partial_x P = \frac{\Delta P}{L} , \tag{3.324}$$

if ΔP is the pressure gradient applied between the ends of the pipe. Transforming the two-dimensional Laplacian in the first equation (3.323) to plane polar coordinates and taking into account that the flow must be symmetric about the symmetry axis of the pipe, we find the equation

$$\frac{\Delta P}{L\eta} = \frac{1}{r}\partial_r\left(r\partial_r v\right) , \tag{3.325}$$

which can easily be integrated to determine the velocity profile

$$v(r) = \frac{\Delta P}{4L\eta}r^2 + A\ln r + B , \tag{3.326}$$

with two integration constants A and B. We must require that $\upsilon(r) = 0$ at the wall of the pipe at $r = R$ and that $\upsilon(r)$ remains regular at $r = 0$. This can be achieved by setting $A = 0$ and

$$B = -\frac{\Delta P}{4L\eta}R^2 , \tag{3.327}$$

which leaves us with the parabolic velocity profile

$$\upsilon(r) = \frac{\Delta P}{4L\eta}\left(r^2 - R^2\right) \tag{3.328}$$

across the pipe (Figure 3.9). The amount of mass flowing through the pipe per unit time is

$$\dot{M} = 2\pi \int_0^R r\mathrm{d}r\,\rho\upsilon(r) = \frac{\pi\Delta P\rho R^4}{8L\eta} , \tag{3.329}$$

which is the Hagen-Poiseulle law: The mass of a viscous fluid flowing through a pipe per unit time is proportional to the squared cross section of the pipe.

Problems

1. A cylinder that contains an incompressible fluid rotates with constant angular velocity $\vec{\omega} = \omega\hat{e}_z$ in the gravitational field of the Earth, characterised by the gravitational acceleration $\vec{g} = -g\hat{e}_z$.

   (a) Use Euler's equation of momentum conservation

   $$\partial_t\vec{\upsilon} + \left(\vec{\upsilon}\cdot\vec{\nabla}\right)\vec{\upsilon} = -\frac{\vec{\nabla}P}{\rho} - \vec{\nabla}\Phi \tag{3.330}$$

   to derive differential equations for each velocity component that contain the angular velocity ω and the gravitational acceleration g. Why does the continuity equation

   $$\partial_t\rho + \vec{\nabla}\cdot(\rho\vec{\upsilon}) = 0 \tag{3.331}$$

   not yield any additional information?

   (b) Determine a function $P(r, \varphi, z)$ from these differential equations. What does the surface of the rotating fluid look like?

2. Jets are large directed outflows of material and a common astrophysical phenomenon. They can be observed under various circumstances, e.g. together with young T Tauri stars and the accretion onto a black hole in the centre of an active galaxy. Here, we want to examine some basic properties of a jet. Assume that a stationary jet has its origin on the surface of a spherical star with mass M and radius R, has initially the velocity υ_0 and the cross-sectional area A_0. The outflowing material has a polytropic equation-of-state, $P = P_0(\rho/\rho_0)^\gamma$, where γ is the adiabatic index, and the entropy stays constant, $\mathrm{d}s = 0$ along flow lines. The gas surrounding the star is assumed to be adiabatic, i.e. the pressure drops exponentially with the distance r from the surface, $P(r) = P_0\exp(-r/h)$, where h is the pressure scale height and P_0 the pressure at the surface.

(a) Determine the specific enthalpy $\tilde{h}$ per unit mass as a function of P and ρ.

(b) Use Bernoulli's equation,

$$\frac{v^2}{2} + \tilde{h} + \Phi = \text{const.} \,, \tag{3.332}$$

along flow lines, to determine $v(r)$.

(c) Use the continuity equation

$$\partial_t \rho + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \tag{3.333}$$

to determine the cross-sectional area $A(r)$.

3. Assume that a layer of height h of a viscous and incompressible fluid flows down a plane inclined by an angle α relative to the horizontal. The top of the fluid is free and feels the atmospheric pressure P_0. The coordinate system is chosen such that the x-axis is parallel to the velocity vector of the fluid and the z-axis is perpendicular to the plane.

(a) Determine the two equations that the Navier-Stokes equation

$$\rho \mathrm{d}_t \vec{v} = -\vec{\nabla} P + \eta \vec{\nabla}^2 \vec{v} + \left(\frac{\eta}{3} + \zeta\right) \vec{\nabla} \left(\vec{\nabla} \cdot \vec{v}\right) + \rho \vec{g} \tag{3.334}$$

simplifies to, where $\vec{g}$ is the gravitational acceleration.

(b) Solve these two differential equations for $P(z)$ and $v(z)$. What are the appropriate boundary conditions to be chosen for $P(z = h)$, $v(z = 0)$ and $(\mathrm{d}v/\mathrm{d}z)(z = h)$?

3.5 Shock waves

This section deals with the formation and the properties of shock waves. We begin with the method of characteristics for quasi-linear systems of partial differential equations and derive the Riemann invariants (3.359), which are used to explain the steepening of non-linear sound waves. Then, we turn to global properties of shock waves following from conservation laws, finding the Rankine-Hugoniot shock jump conditions (3.376) and (3.377). The velocity (3.388) of the shock itself relative to the flow is derived and used in the derivation of Sedov's solution (3.399) for the outer radius of a strong spherical shock wave.

The hydrodynamical equations are a set of non-linear, partial differential equations which give rise to non-linear phenomena in fluid flows. One important aspect is the formation of shock waves, where the velocity field changes discontinuously. Despite the non-linearity of the equations and some of the phenomena they describe, some statements can be made on characteristic properties of the flow without even solving the hydrodynamical equations. We have seen some examples before, such as Kelvin's circulation theorem and Bernoulli's law. We shall now proceed to show that even strongly non-linear phenomena such as shock waves can be predicted as inevitable, and that some important properties they display can be generally given. For doing so, we restrict our treatment to one-dimensional flows, having in mind fluid flows in pipes, for example. We shall begin with the method of characteristics.

3.5.1 The method of characteristics

In one dimension, for a polytropic, inviscid fluid, the continuity and Euler equations simplify to

$$\begin{aligned} \partial_t\rho + \rho\partial_x v + v\partial_x\rho &= 0 \ , \\ \partial_t v + v\partial_x v + \frac{\partial_x P}{\rho} &= 0 \ . \end{aligned} \tag{3.335}$$

The derivative of the pressure can be expressed by the derivative of the density,

$$\partial_x P = c_s^2 \partial_x \rho \ , \tag{3.336}$$

introducing the sound speed

$$c_s^2 = c_{s0}^2 \left(\frac{\rho}{\rho_0}\right)^{\gamma-1} \ . \tag{3.337}$$

Taking the differential of the last equation, we see that the differentials of the sound velocity and of the density are related by

$$2\frac{\mathrm{d}c_s}{c_s} = (\gamma-1)\frac{\mathrm{d}\rho}{\rho} \ , \tag{3.338}$$

which enables us to replace the partial density derivates of the density according to

$$\frac{\partial_t\rho}{\rho} = \frac{2}{\gamma-1}\frac{\partial_t c_s}{c_s} \ , \quad \frac{\partial_x\rho}{\rho} = \frac{2}{\gamma-1}\frac{\partial_x c_s}{c_s} \ . \tag{3.339}$$

Our reduced set of one-dimensional hydrodynamical equations now reads

$$\begin{aligned} \frac{2}{\gamma-1}\partial_t c_s + c_s\partial_x v + \frac{2v}{\gamma-1}\partial_x c_s &= 0 \ , \\ \partial_t v + v\partial_x v + \frac{2c_s}{\gamma-1}\partial_x c_s &= 0 \ . \end{aligned} \tag{3.340}$$

They are two partial differential equations for two functions, c_s and v, in two variables, t and x.

It is important to see that these equations are quasi-linear, which means that the highest-order derivatives of the unknown functions occur linearly in them. Due to this property, we can summarise the two equations as

$$\begin{pmatrix} \frac{2}{\gamma-1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \partial_t c_s \\ \partial_t v \end{pmatrix} + \begin{pmatrix} \frac{2v}{\gamma-1} & c_s \\ \frac{2c_s}{\gamma-1} & v \end{pmatrix}\begin{pmatrix} \partial_x c_s \\ \partial_x v \end{pmatrix} = 0 \ . \tag{3.341}$$

At this point, the method of characteristics sets in.

Suppose, more generally, that we are given a set of n quasi-linear, partial differential equations for the n unknown functions u_j of the two variables x and y. By its quasi-linearity, this set of equations can be brought into the form

$$X_{ij}\partial_x u_j + Y_{ij}\partial_y u_j = Z_i \ , \tag{3.342}$$

where the Z_i represent possible inhomogeneities of the equations. The method of characteristics consists in finding local directions in the x-y plane into which the partial differential equations can be written as complete differentials, and thus be integrated. Of course, the functions could also depend on more than two independent variables, but we restrict the discussion to this case here for simplicity.

We wish to find differentials $\mathrm{d}\vec{s}$ in the two-dimensional space of independent variables satisfying the conditions

$$\mathrm{d}\vec{s}^{\top} X = \vec{L}^{\top}\mathrm{d}x\,, \quad \mathrm{d}\vec{s}^{\top} Y = \vec{L}^{\top}\mathrm{d}y \tag{3.343}$$

with the same vector $\vec{L}$ on the right-hand sides, for both matrices X and Y. If we could find such differentials, with a vector $\vec{L}$ yet to be determined, multiplying our set of equations (3.342) with it from the left would result in

$$L_j\left(\partial_x u_j \mathrm{d}x + \partial_y u_j \mathrm{d}y\right) = L_j \mathrm{d}u_j = \mathrm{d}s_i Z_i\,. \tag{3.344}$$

We could then directly integrate these equations, finding

$$L_j u_j = Z_i s_i\,. \tag{3.345}$$

In order to see when we can hope to find such a vector of differentials $\mathrm{d}\vec{s}$, we multiply the first equation (3.343) by $\mathrm{d}y$, the second by $\mathrm{d}x$ and subtract the second from the first to get

$$\mathrm{d}\vec{s}^{\top}\left(X\mathrm{d}y - Y\mathrm{d}x\right) = 0\,. \tag{3.346}$$

For this set of linear equations to have a non-trivial solution for $\mathrm{d}\vec{s}$, the determinant of the matrix $X\mathrm{d}y - Y\mathrm{d}x$ must vanish,

$$\det\left(X\mathrm{d}y - Y\mathrm{d}x\right) = 0\,. \tag{3.347}$$

This will give us relations between the two differentials $\mathrm{d}y$ and $\mathrm{d}x$ which will define preferred directions in x-y space. The integral curves of the expressions for $\mathrm{d}y/\mathrm{d}x$ are the *characteristics* of the system (3.342) of quasi-linear partial differential equations. The differentials are then found as eigenvectors of the matrix $X\mathrm{d}y - Y\mathrm{d}x$ belonging to the eigenvalue zero. Once they have been found, the vector $\vec{L}$ is given by the two equations (3.343).

Let us apply this method of characteristics now to the set of hydrodynamical equations (3.341). Here, we have the two functions c_s and v in place of the u_1 and u_2, and the two independent variables (t, x) in place of (x, y). The two matrices X and Y are replaced by

$$T = \begin{pmatrix} \frac{2}{\gamma-1} & 0 \\ 0 & 1 \end{pmatrix}\,, \quad X = \begin{pmatrix} \frac{2v}{\gamma-1} & c_\mathrm{s} \\ \frac{2c_\mathrm{s}}{\gamma-1} & v \end{pmatrix}\,. \tag{3.348}$$

The characteristics are defined by the condition

$$\begin{aligned} 0 = \det\left(T\mathrm{d}x - X\mathrm{d}t\right) &= \det\begin{pmatrix} \frac{2(\mathrm{d}x - v\mathrm{d}t)}{\gamma-1} & -c_\mathrm{s}\mathrm{d}t \\ -\frac{2c_\mathrm{s}\mathrm{d}t}{\gamma-1} & \mathrm{d}x - v\mathrm{d}t \end{pmatrix} \\ &= \frac{2}{\gamma-1}\det\begin{pmatrix} \mathrm{d}x - v\mathrm{d}t & -c_\mathrm{s}\mathrm{d}t \\ -c_\mathrm{s}\mathrm{d}t & \mathrm{d}x - v\mathrm{d}t \end{pmatrix}\,, \end{aligned} \tag{3.349}$$

which leads to the quadratic equation

$$\mathrm{d}x^2 - 2v\mathrm{d}x\mathrm{d}t + \left(v^2 - c_\mathrm{s}^2\right)\mathrm{d}t^2 = 0\;, \tag{3.350}$$

whose two solutions

$$\mathrm{d}x_\pm = (v \pm c_\mathrm{s})\mathrm{d}t \tag{3.351}$$

define the characteristics of the system (3.341) of hydrodynamical equations.

?

What does the condition (3.351) mean geometrically in a space-time diagram?

The differentials ds must be eigenvectors with eigenvalue zero of the difference matrix $T\mathrm{d}x - X\mathrm{d}t$,

$$(\mathrm{d}s_1, \mathrm{d}s_2)\cdot\begin{pmatrix} \frac{2(\mathrm{d}x-v\mathrm{d}t)}{\gamma-1} & -c_\mathrm{s}\mathrm{d}t \\ -\frac{2c_\mathrm{s}\mathrm{d}t}{\gamma-1} & \mathrm{d}x - v\mathrm{d}t \end{pmatrix} = (0,0)\;. \tag{3.352}$$

In particular, this establishes the relation

$$-c_\mathrm{s}\mathrm{d}t\mathrm{d}s_1 + (\mathrm{d}x - v\mathrm{d}t)\mathrm{d}s_2 = 0 \tag{3.353}$$

between ds_1 and ds_2. On the characteristics, $\mathrm{d}x = \mathrm{d}x_\pm = (v \pm c_\mathrm{s})\mathrm{d}t$, hence d$s_1$ and ds_2 must agree except for their sign,

$$-c_\mathrm{s}\mathrm{d}s_1 \pm c_\mathrm{s}\mathrm{d}s_2 = 0 \quad\Rightarrow\quad \mathrm{d}s_2 = \pm\mathrm{d}s_1\;. \tag{3.354}$$

The vector $\vec{L}$ is finally found from one of the equations (3.343) applied to our current situation,

$$(\mathrm{d}s_1, \pm\mathrm{d}s_1)\begin{pmatrix} \frac{2v}{\gamma-1} & c_\mathrm{s} \\ \frac{2c_\mathrm{s}}{\gamma-1} & v \end{pmatrix} = (L_1, L_2)\mathrm{d}x\;, \tag{3.355}$$

which implies

$$L_1\mathrm{d}x = \frac{2(v \pm c_\mathrm{s})}{\gamma - 1}\mathrm{d}s_1\;,\quad L_2\mathrm{d}x = (c_\mathrm{s} \pm v)\mathrm{d}s_1\;. \tag{3.356}$$

The ratio between these two components is all we need because d$\vec{s}$ and d$\vec{L}$ can only be determined up to a common normalisation factor. The last equation tells us

$$\frac{L_1}{L_2} = \pm\frac{2}{\gamma - 1}\;. \tag{3.357}$$

We arbitrarily set $L_2 = 1$ and return to evaluate (3.344) for our hydrodynamical equations, where the inhomogeneities $Z_i = 0$. This finally leads us to

$$\mathrm{d}v \pm \frac{2}{\gamma - 1}\mathrm{d}c_\mathrm{s} = 0\;, \tag{3.358}$$

which we can directly integrate to find the two *Riemann invariants*

$$R_\pm = v \pm \frac{2}{\gamma - 1}c_\mathrm{s}\;, \tag{3.359}$$

which are conserved on the plus and minus characteristics, respectively.

3.5.2 Steepening of sound waves

Consider now a hitherto unperturbed fluid at rest, on which a non-linear density perturbation is imprinted at time $t = 0$. Every point within the perturbation can be connected to its unperturbed neighbourhood by means of minus characteristics coming from the positive x region in the past. Those characteristics are straight lines with slope

$$\frac{\mathrm{d}x_-}{\mathrm{d}t} = v - c_\mathrm{s} = -c_{\mathrm{s}0} \tag{3.360}$$

because they propagate through unperturbed material at rest. Along these minus characteristics, the Riemann invariant

$$R_- = v - \frac{2c_\mathrm{s}}{\gamma - 1} = -\frac{2c_{\mathrm{s}0}}{\gamma - 1} \tag{3.361}$$

is conserved. This establishes the relation

$$v = \frac{2(c_\mathrm{s} - c_{\mathrm{s}0})}{\gamma - 1} \tag{3.362}$$

at every point that can be reached by a minus characteristic coming from unperturbed material, which is every point in the fluid.

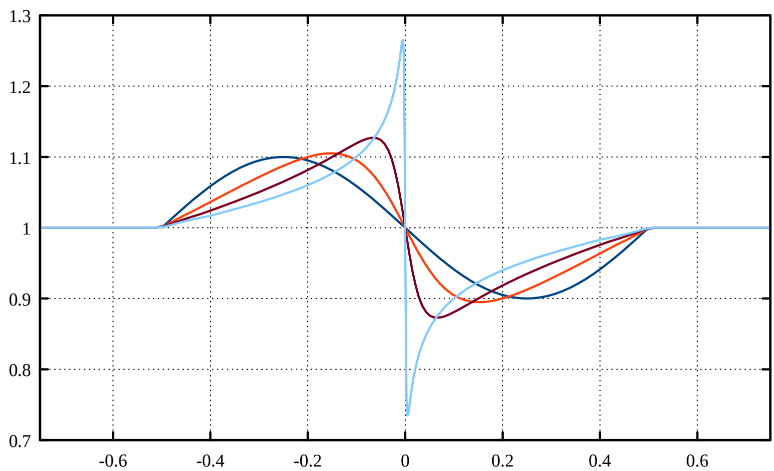

Figure 3.10 Illustration of how a non-linear density wave steepens and ultimately turns into a shock.

At the boundary points of the perturbation, the density is supposed to have dropped to the unperturbed density, and they are considered to be at rest. A plus characteristic attached to the boundary point at $x > 0$ is determined by

$$\frac{\mathrm{d}x_+}{\mathrm{d}t} = v + c_\mathrm{s} = c_{\mathrm{s}0}\ , \tag{3.363}$$

because it propagates into unperturbed material. Along this plus characteristic, the Riemann invariant

$$R_+ = v + \frac{2c_\mathrm{s}}{\gamma - 1} = \frac{2c_{\mathrm{s}0}}{\gamma - 1} \tag{3.364}$$

is conserved. The right boundary point of the perturbation has unperturbed density and retains it as it propagates along the plus characteristic. Now consider the central point of the density perturbation, where the density is highest. Its enhanced density implies a sound speed higher than that of the unperturbed fluid. According to (3.362), overdense points have a velocity $v > 0$. Both, the velocity v and the sound speed c_s, are therefore higher at an overdensity than in the unperturbed fluid. Plus characteristics originating there thus propagate faster than plus characteristics originating from unperturbed points. The plus characteristic of the density peak thus approaches that of the right boundary point of the perturbation. The density peak will catch up with the boundary point and ultimately reach it: The density perturbation steepens and ultimately produces a discontinuity in the density and the velocity because streams of different density and velocity cannot coexist at the same location in a fluid (Figure 3.10).

We have thus shown simply by the method of characteristics that non-linear density perturbations have to steepen and ultimately form discontinuities, or shocks. As generic as our discussion was, as generic is this result: The formation of shocks by steepening of non-linear waves is inevitable in a fluid. It is quite remarkable that we did not have to solve any of the hydrodynamical equations to see this. The method of characteristics was sufficient.

3.5.3 The Rankine-Hugoniot shock jump conditions

Even though at least some of the flow variables may be discontinuous at a shock, three current densities must be conserved across the shock, namely the matter current density $\rho\vec{v}$, the energy current density

$$\vec{q} = \left(\frac{v^2}{2} + \tilde{h}\right)\rho\vec{v} \tag{3.365}$$

and the momentum-current density

$$T^{ij} = \rho v^i v^j + P\delta^{ij} \; . \tag{3.366}$$

We consider now a shock that is perpendicular to the local flow direction. We fix an arbitrary point on the shock surface and locally construct a coordinate system such that the x axis is perpendicular to the shock surface and the y-z plane is tangential to it. On the y-z plane, the three conserved current densities must meet. Identifying with subscripts 1 and 2 quantities on either side of the shock, we must have

$$\begin{aligned}
\rho_1 v_1 &= \rho_2 v_2 \; , \\
\left(\frac{v_1^2}{2} + \tilde{h}_1\right)\rho_1 v_1 &= \left(\frac{v_2^2}{2} + \tilde{h}_2\right)\rho_2 v_2 \; , \\
\rho_1 v_1^2 + P_1 &= \rho_2 v_2^2 + P_2 \; .
\end{aligned} \tag{3.367}$$

We wish to express the flow variables ρ_2, v_2 and P_2 on one side of the shock by those on the other. For doing so, we first adopt a polytropic equation of state and thereby fix the sound speed

$$c_\mathrm{s}^2 = \gamma\frac{P}{\rho} \tag{3.368}$$

and the specific enthalpy per unit mass

$$\tilde{h} = \frac{\gamma}{\gamma - 1}\frac{P}{\rho} = \frac{c_\mathrm{s}^2}{\gamma - 1} \,. \tag{3.369}$$

We further introduce the Mach number on the 1-side of the shock,

$$v_1^2 = \mathcal{M}_1^2 c_{\mathrm{s}1}^2 \,, \tag{3.370}$$

and the ratios r and q between the density and the pressure values on both sides of the shock,

$$r := \frac{\rho_2}{\rho_1} \,, \quad q := \frac{P_2}{P_1} \,. \tag{3.371}$$

The enthalpy on the 2-side of the shock is then

$$\tilde{h}_2 = \frac{\gamma}{\gamma - 1}\frac{P_2}{\rho_2} = \frac{\gamma}{\gamma - 1}\frac{P_1}{\rho_1}\frac{P_2}{P_1}\frac{\rho_1}{\rho_2} = \tilde{h}_1\frac{q}{r} = \frac{c_{\mathrm{s}1}^2}{\gamma - 1}\frac{q}{r} \,. \tag{3.372}$$

After this preparation, (3.367) can be reduced to

$$\begin{aligned} \frac{\mathcal{M}_1^2}{2} + \frac{1}{\gamma - 1} &= \frac{\mathcal{M}_1^2}{2r^2} + \frac{1}{\gamma - 1}\frac{q}{r} \,, \\ \mathcal{M}_1^2 + \frac{1}{\gamma} &= \frac{\mathcal{M}_1^2}{r} + \frac{q}{\gamma} \,. \end{aligned} \tag{3.373}$$

We multiply the first of these equations with $r(\gamma - 1)$ and the second with γ to find

$$\begin{aligned} \frac{\mathcal{M}_1^2}{2} r(\gamma - 1) + r &= \frac{\mathcal{M}_1^2}{2r}(\gamma - 1) + q \,, \\ \mathcal{M}_1^2\gamma + 1 &= \frac{\mathcal{M}_1^2}{r}\gamma + q \,, \end{aligned} \tag{3.374}$$

and subtract the first from the second to elimitate q and retain the quadratic equation in r

$$r^2\left[\mathcal{M}_1^2(\gamma - 1) + 2\right] - 2r\left(\mathcal{M}_1^2\gamma + 1\right) + \mathcal{M}_1^2(\gamma + 1) = 0 \,, \tag{3.375}$$

which has the two solutions

$$r_+ =: \boxed{r = \frac{\mathcal{M}_1^2(\gamma + 1)}{\mathcal{M}_1^2(\gamma - 1) + 2}} \,, \quad r_- = 1 \,. \tag{3.376}$$

Only the solution r_+ is interesting since r_- corresponds to no density jump at all. We thus set $r = r_+$ and use this to find q from the second equation (3.374),

$$\boxed{q = \frac{2\gamma\mathcal{M}_1^2 - \gamma + 1}{\gamma + 1}} \,. \tag{3.377}$$

?

Repeat the derivation of the jump conditions (3.376) and (3.377) on your own.

The temperature jump can finally be obtained from the equation of state, such as the ideal-gas equation, through

$$\frac{T_2}{T_1} = \frac{P_2}{\rho_2}\frac{\rho_1}{P_1} = \frac{q}{r} \,. \tag{3.378}$$

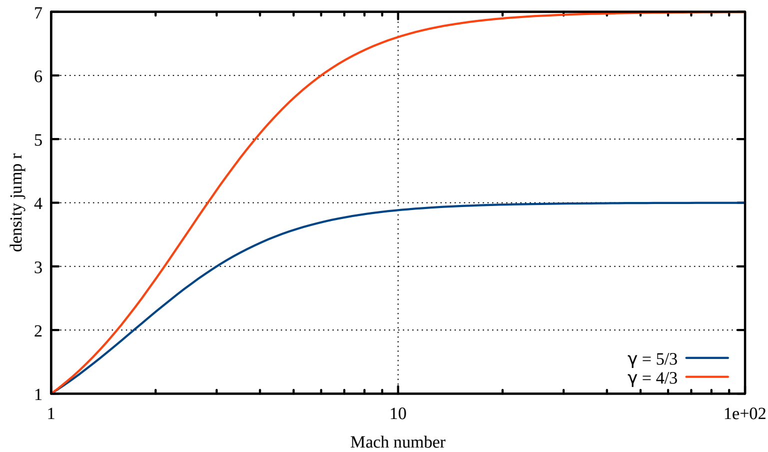

Figure 3.11 On the Rankine-Hugoniot shock jump conditions: The density jump r at a shock is shown as a function of the Mach number upstream the shock for gases with adiabatic indices $\gamma = 5/3$ and $\gamma = 4/3$.

The gas is assumed to approach the shock with supersonic velocity, $\mathcal{M}_1 > 1$. Then, (3.376) tells us that

$$r = \frac{\gamma+1}{\gamma-1+2\mathcal{M}_1^{-2}} = \frac{\gamma+1}{\gamma+1+2(\mathcal{M}_1^{-2}-1)} > 1 \,. \tag{3.379}$$

The density is higher downstream of the shock, the velocity must correspondingly be lower. According to (3.377), the pressure also increases because

$$q = \frac{2\gamma\mathcal{M}_1^2-\gamma+1}{\gamma+1} > \frac{2\gamma-\gamma+1}{\gamma+1} = 1 \,. \tag{3.380}$$

In the ultrasonic limit, $\mathcal{M}_1 \gg 1$, the density jump approaches

$$r = \frac{\gamma+1}{\gamma-1} \,, \tag{3.381}$$

or $r = 4$ for a monatomic, ideal, non-relativistic gas. The closer γ gets towards unity, the larger r will become (Figure 3.11). For a relativistic gas, $\gamma = 4/3$ and $r = 7$. The pressure and temperature jumps across such strong shocks can become arbitrarily large. Both can rise strongly, showing that the gas will be hot downstream the shock.

Caution The adiabatic index for a gas composed of molecules with f degrees of freedom is

$$\gamma = \frac{f+2}{f}$$

for a non-relativistic and

$$\gamma = \frac{f+1}{f}$$

for a relativistic gas. For $f = 3$, i.e. if the molecules have only the three translational but no internal (rotational or vibrational) degrees of freedom, $\gamma = 5/3$ in the non-relativistic and $\gamma = 4/3$ in the relativistic case. ◀

Eliminating the Mach number between (3.376) and (3.377) gives either of the two equations

$$r = \frac{(\gamma+1)q+(\gamma-1)}{(\gamma-1)q+(\gamma+1)} \,, \quad q = \frac{(\gamma+1)r-(\gamma-1)}{(\gamma+1)-(\gamma-1)r} \tag{3.382}$$

relating the pressure and the density jumps to each other.

3.5.4 Shock velocity

Having seen how conservation laws alone predict the discontinuities in the density, the velocity and the pressure of a fluid through the Rankine-Hugoniot conditions, we now turn to the question how fast the shock itself will propagate through the fluid. To this end, consider a long pipe filled with gas and closed with a piston at its left end. At time $t = 0$, we imagine that the piston is instantaneously accelerated to some high and constant velocity u.

The sudden acceleration will drive a shock into the unperturbed gas ahead of the piston. Downstream of the shock, the gas will be at rest, $v_1 = 0$, while it will have the velocity of the piston, $v_2 = u$, upstream. In between, the shock will move with a yet unknown velocity v_s.

To analyse this situation, we transform into the rest frame of the shock and mark all velocities in the rest frame of the shock with primes. In that frame, by construction, $v'_\mathrm{s} = 0$, further $v'_1 = v_1 - v_\mathrm{s} = -v_\mathrm{s}$ and $v'_2 = v_2 - v_\mathrm{s} = u - v_\mathrm{s}$, while the velocity difference between down- and upstream remains of course unchanged, $v_2 - v_1 = u = v'_2 - v'_1$.

Solving the identity

$$u = v'_2 - v'_1 = v'_1 \left(\frac{1}{r} - 1\right) \tag{3.383}$$

for $v'_1 = -v_\mathrm{s}$ immediately gives the shock velocity

$$v_\mathrm{s} = \frac{ru}{r-1} \tag{3.384}$$

in terms of the velocity u of the piston. A strong shock, which has $r = (\gamma + 1)/(\gamma - 1)$ as we have seen before, thus moves with the velocity

$$v_\mathrm{s} = \frac{\gamma + 1}{2}\, u\ . \tag{3.385}$$

?

Test several limiting cases of the result (3.388) for the shock velocity and interpret the results.

In an ideal, monatomic, nonrelativistic gas, for example, the shock velocity exceeds the velocity of the piston by $4/3 - 1 \approx 33\,\%$. We now know the shock speed only as a function of the velocity u of the piston. Sometimes this is unknown, sometimes it is irrelevant because we want to know the shock speed in terms of the intrinsic properties r and q of the shock. To achieve this, we simply write

$$v_\mathrm{s} = -v'_1 = |\mathcal{M}_1| c_{\mathrm{s}1}\ , \tag{3.386}$$

solve either one of the Rankine-Hugoniot shock jump conditions (3.376) or (3.377) for $\mathcal{M}_1^2$,

$$\mathcal{M}_1^2 = \frac{2r}{(\gamma+1) - (\gamma-1)r} = \frac{(\gamma+1)q + (\gamma-1)}{2\gamma}\ , \tag{3.387}$$

and insert the result into (3.386) to find the shock velocity as a function of either the density jump r or the pressure jump q,

$$v_\mathrm{s} = c_{\mathrm{s}1}\sqrt{\frac{2r}{(\gamma+1) - (\gamma-1)r}} = c_{\mathrm{s}1}\sqrt{\frac{(\gamma+1)q + (\gamma-1)}{2\gamma}}\ . \tag{3.388}$$

3.5.5 The Sedov solution

An impressive example for a shock wave is given by an explosion, i.e. by an event in which in very short time energy is being released within a very small volume (Figure 3.12). We now consider such an event under the following simplifying assumptions: (1) The shock is very strong, meaning that the pressure of the surrounding medium can be neglected, $P_1 \ll P_2$. (2) The explosion energy E is released instantaneously and the energy of the surrounding material is negligible compared to E, i.e. the explosion energy dominates that of the surroundings. (3) The gas be polytropic with an adiabatic index γ.

Under these conditions, our shock jump condition for the density is

$$r = \frac{\rho_2}{\rho_1} \approx \frac{\gamma + 1}{\gamma - 1} , \tag{3.389}$$

as it has to be for a very strong shock. The densities ρ_1 and ρ_2 are completely determined by each other, which implies that the behaviour of the shock must be entirely determined by the explosion energy E and the surrounding matter density ρ_1.

Let us now consider the shock at a time t after the explosion, when it has already reached an unknown radius $R(t)$. The only quantity with the dimension of a length that can be formed from E, t and ρ_1 is

$$\left(\frac{Et^2}{\rho_1}\right)^{1/5} , \tag{3.390}$$

which suggests the ansatz

$$R(t) = R_0 \left(\frac{Et^2}{\rho_1}\right)^{1/5} \tag{3.391}$$

with a dimension-less constant R_0 which remains to be determined. The shock velocity is obviously the time derivative of $R(t)$,

$$v_\mathrm{s} = \frac{\mathrm{d}R}{\mathrm{d}t} = \frac{2}{5}\frac{R}{t} = \frac{2R_0}{5}\left(\frac{E}{\rho_1}\right)^{1/5} t^{-3/5} , \tag{3.392}$$

but it also has to obey the relation (3.388) found above. Solving the latter for the pressure jump q and inserting $c_\mathrm{s}^2 = \gamma P_1/\rho_1$ gives

$$q = \frac{P_2}{P_1} = \frac{1}{\gamma + 1}\left[\frac{2v_\mathrm{s}^2\rho_1}{P_1} - (\gamma - 1)\right] . \tag{3.393}$$

Under the assumption that $P_1 \ll P_2$, we can neglect the second term on the right-hand side and approximate the pressure inside the shock by

$$P_2 = \frac{2v_\mathrm{s}^2\rho_1}{\gamma + 1} . \tag{3.394}$$

The jump condition (3.389) for the density ratio shows that the density inside the shock must remains constant in time because ρ_1 is constant. Given the shock velocity (3.392), the pressure inside the shock must be

$$P_2 = \frac{8R_0^2}{25(\gamma + 1)} E^{2/5}\rho_1^{3/5}t^{-6/5} . \tag{3.395}$$

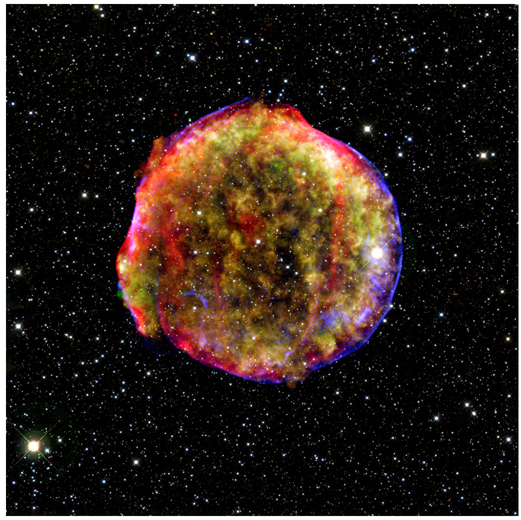

Figure 3.12 This image of the remnant of Tycho's supernova combines X-ray and infrared data from the Chandra and Spitzer space telescopes, respectively. Blue-red colours show the energetic X-ray emission at the shock, while yellow-green colours show the infrared emission inside the shock front.

According to (3.385), the velocity of the gas behind the shock then falls off in the same way,

$$u = \frac{2v_s}{\gamma + 1} = \frac{4R_0}{5(\gamma + 1)} \left(\frac{E}{\rho_1}\right)^{1/5} t^{-3/5} \,. \tag{3.396}$$

We can interpret these relations as follows: A shock wave driven by the sudden release of a large amount of energy E, which propagates outward with the time-dependent radius $R(t)$, sweeps up surrounding material with the mass

$$M = \frac{4\pi}{3}\rho_1 R^3 \,, \tag{3.397}$$

which is accelerated from rest to a velocity $\approx R/t$. Thus, the kinetic energy

$$E_{\text{kin}} \approx \frac{4\pi}{3}\rho_1 \frac{R^5}{t^2} \tag{3.398}$$

must be put into the swept-up material. Equating this to the explosion energy E, we immediately find

$$R = \left(\frac{3Et^2}{4\pi\rho_1}\right)^{1/5} \,, \tag{3.399}$$

i.e. the scaling relation (3.391) simply expresses energy conservation.

Without solving any of the hydrodynamical equations, we now know how the radius of the explosion shock, its velocity as well as the pressure and the density at its inside. They are completely determined by the released amount of energy E and the density ρ_1 of the surrounding material.

Problems

1. In astrophysics, the hydrodynamical equations often simplify a lot due to spherical symmetry, e.g. when a young star of mass M accretes material. In this case, $\vec{v} = v(r)\hat{e}_r$, and the continuity and the Euler equations simplify to

$$\partial_t \rho + \frac{1}{r^2}\partial_r\left(r^2 \rho v\right) = 0 \ , \quad \partial_t v + v\partial_r v + \frac{c_s^2}{\rho}\partial_r \rho = -\partial_r \Phi = \frac{GM}{r^2} \ , \quad (3.400)$$

respectively, where we have used $\vec{\nabla} P = c_s^2 \vec{\nabla}\rho$. Let us for simplicity further assume that the sound speed c_s is constant. In order to use the method of characteristics, the equations have to be brought into the form

$$T_{ij}\partial_t u^j + R_{ij}\partial_r u^j = Z_i \ , \quad (3.401)$$

where summing over j is implied. The matrix elements R_{ij} and T_{ij} are given by the coefficients in front of the partial derivatives, while the Z_i are given by the inhomogeneities.

   (a) Bring the continuity and the Euler equations into the form (3.401) and identify T, R, $\vec{u}$ and $\vec{Z}$.

   (b) Determine a relation between the differentials dr and dt from the condition $\det(T\mathrm{d}r - R\mathrm{d}t) = 0$. What does the result mean physically?

   (c) The goal is to find the differentials $\mathrm{d}s^\top = (\mathrm{d}s_1, \mathrm{d}s_2)$ and the vector $\vec{L}^\top = (L_1, L_2)$ such that

$$\mathrm{d}\vec{s}^\top T = \vec{L}^\top \mathrm{d}t \ , \quad \mathrm{d}\vec{s}^\top T = \vec{L}^\top \mathrm{d}r \quad (3.402)$$

   are satisfied. Determine ds_1 and ds_2 from $\mathrm{d}\vec{s}^\top(T\mathrm{d}r - R\mathrm{d}t) = 0$ and L_1 from (3.402), arbitrarily setting $L_2 = 1$.

   (d) Multiplying (3.401) by $\mathrm{d}\vec{s}^\top$ from the left and using (3.402) leads to $\vec{L}^\top \cdot \mathrm{d}\vec{u} = \mathrm{d}\vec{s}^\top \cdot \vec{Z}$. Set up the latter equation which defines the Riemann invariants and carry the necessary integration out as far as possible.

2. Assume that the coordinate system is chosen such that the yz-plane is parallel to a shock front and the x-direction perpendicular to it and the gas flows from the side 1 to the side 2. With the energy-momentum tensor

$$T^{\mu\nu} = \left(\rho + \frac{P}{c^2}\right)u^\mu u^\nu - \eta^{\mu\nu}P \quad (3.403)$$

and the four-velocity $(u^\mu) = \gamma(c, \vec{v}^\top)$, where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the Lorentz factor, the relativistic generalisations of the continuity conditions for the densities of the particle current, the momentum current and the energy current are given by

$$n_1 u_1^x = n_1 u_1^x \ , \quad T_1^{xx} = T_2^{xx} \ , \quad cT_1^{0x} = cT_2^{0x} \ , \quad (3.404)$$

respectively.

(a) Express the continuity conditions as functions of the velocities $\beta_i \equiv \beta_i^x$ in units of the light speed, the enthalpies per volume $h_i = \varepsilon_i + P_i$ with $\varepsilon_i = \rho c^2$, and the specific volumes per particle $V_i \equiv n_i^{-1}$ (here and in the following, $i = 1, 2$).

(b) Determine the velocities on both sides of the discontinuity as a function of P_i and ε_i. *Hint*: It is helpful to introduce the rapidity $\theta \equiv \operatorname{artanh}(v/c)$.

(c) What is the relative velocity v_{12} of the gases on both sides of the discontinuity?

(d) In the ultrarelativistic case, $P = \varepsilon/3$. Determine the velocities on both sides of the discontinuity in the case of a very strong shock front ($\varepsilon_2 \to \infty$).

3.6 Instabilities

This section concludes the chapter on hydrodynamics with a discussion of fluid instabilities. The method of analysis is common to most of them: The governing equations are linearised by a perturbation ansatz. Decomposing the perturbations into plane waves turns these linearised differential equations into systems of linear algebraic equations which the dispersion relations can directly be derived from requiring non-trivial solutions. We begin with surface waves on a fluid in a gravitational field which are shown to satisfy the non-linear dispersion relation (3.418). The Rayleigh-Taylor or buoyancy instability follows, whose dispersion relation (3.426) shows that a specifically lighter fluid placed below a specifically heavier fluid tends to develop an unstable boundary. The Kelvin-Helmholtz or shear instability arises at the boundary between two fluids one of which flows with respect to the other. Its dispersion relation is given in (3.436). Thermal instability sets in if and when heating a gas leads to less efficient cooling, or cooling leads to less efficient heating. The dispersion relation of this instability is shown in (3.456). After a brief intermezzo on heat conduction, we consider heat transport by convection and show in (3.486) that convection sets in if the temperature gradient is steep. Finally, we briefly discuss turbulence and derive the Kolmogorov spectrum (3.500) for the energy distribution over scales of turbulent eddies.

We shall now proceed to examine hydrodynamical instabilities, i.e. the evolution of situations in which an equilibrium configuration is slightly perturbed. Such investigations follow standard procedures. The equilibrium configuration is taken as given. Small perturbations are applied and the relevant equations are linearised in these perturbations. The linear equations resulting therefrom are then decomposed into Fourier modes whose dispersion relation is derived. Instable situations are characterised by complex or imaginary frequencies, which signal exponential growth of the perturbations. For simplicity, we shall assume that the fluids are inviscid and incompressible, hence $\vec{\nabla} \cdot \vec{v} = 0$.

We transform into the rest frame of one of the unperturbed solutions, in such a way that the velocity field there is given by the velocity perturbation only. Then, only terms linear in the velocity need to be retained. The vorticity equation (3.272) then tells us that

$$\partial_t\vec{\Omega} = \partial_t\left(\vec{\nabla}\times\vec{v}\right) = 0\ . \tag{3.405}$$

Vorticity cannot build up, and we can assume that $\vec{\nabla}\times\vec{v} = 0$. Then, there exists a velocity potential ψ such that $\vec{v} = \vec{\nabla}\psi$. Since the fluid is also incompressible, the velocity potential must satisfy Laplace's equation

$$\vec{\nabla}^2\psi = 0\ . \tag{3.406}$$

In addition, Bernoulli's law in the form (3.315) must hold with the term quadratic in the velocity neglected,

$$\partial_t\psi + \tilde{h} + \Phi = 0\ . \tag{3.407}$$

Moreover, for an incompressible fluid, the enthalpy per unit mass is determined by

$$\mathrm{d}\tilde{h} = \frac{\mathrm{d}P}{\rho} = \mathrm{d}\left(\frac{P}{\rho}\right) \tag{3.408}$$

since $\mathrm{d}\rho = 0$, hence $\tilde{h} = P/\rho$, and Bernoulli's law turns into

$$P = -\rho\left(\partial_t\psi + \Phi\right)\ . \tag{3.409}$$

?

Interpret the physical meaning of (3.409). To do so, taking the gradient is helpful.

We begin with two situations in which two fluids are separated by a surface. Our fundamental set of equations for these investigations will be (3.406) and (3.409).

3.6.1 Gravity waves

Consider a fluid which rests under local gravity such that its surface is a plane. Above the surface, we imagine a gas which is much less dense than the fluid and sets the pressure P_2 = const. at the surface. We introduce a coordinate frame such that the surface of the fluid coincides with the x-y plane. We wish to find out how perturbations in the fluid surface propagate.

To this end, we introduce a function $\zeta(x, y, t)$ describing the perturbed fluid surface (Figure 3.13). The velocity in z direction is the change of ζ with time, hence

$$v_z = \partial_t\zeta + \vec{v}\cdot\vec{\nabla}\zeta \approx \partial_t\zeta\ , \tag{3.410}$$

where the last step was possible since $\vec{v}$ and ζ are both small quantities. Moreover, the velocity in z direction is the z derivative of the velocity potential ψ,

$$v_z = \partial_z\psi\ . \tag{3.411}$$

Given the external, e.g. atmospheric pressure P_0, equation (3.409) demands

$$P_0 = -\rho\left(\partial_t\psi + g\zeta\right)\ , \tag{3.412}$$

where $\Phi = g\zeta$ is the local gravitational potential due to the gravitational acceleration g, evaluated at the surface where $z = \zeta$. Since only the spatial

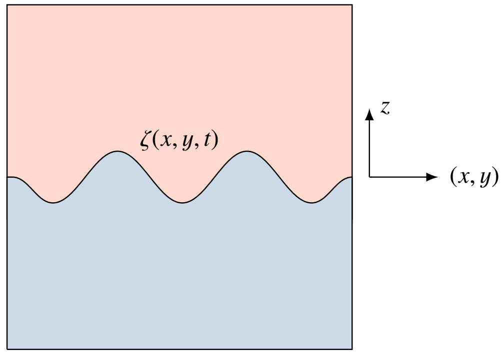

Figure 3.13 Illustration of the perturbed boundary between two fluids, described by a function $z = \zeta(x, y, t)$.

gradient of ψ is relevant for the velocity $\vec{v}$, we can absorb the constant pressure term into ψ by the gauge choice

$$\psi \to \psi + \frac{P_0 t}{\rho} \ . \tag{3.413}$$

The equation we have to solve is then simply

$$\partial_t \psi = -g\zeta \ . \tag{3.414}$$

Taking another time derivative leads to the equation

$$\partial_t \zeta = v_z = \partial_z \psi = -\frac{1}{g}\partial_t^2 \psi \ , \tag{3.415}$$

which has to be evaluated at $z = \zeta$. The velocity potential thus has to satisfy the Laplace equation (3.406) and Bernoulli's equation (3.415).

We begin our solution with the *ansatz*

$$\psi = f(z)\mathrm{e}^{\mathrm{i}(kx - \omega t)} \tag{3.416}$$

for a wave-like solution propagating in x direction. The Laplace equation, applied to this *ansatz*, constrains $f(z)$ to satisfy

$$f''(z) - k^2 f(z) = 0 \ , \tag{3.417}$$

which is solved by $f(z) = f_0 \exp(\pm kz)$. Since this velocity potential is confined to $z < 0$, we need the branch $f(z) = f_0 \exp(kz)$ here.

?

The choice $f(z) = f_0 \mathrm{e}^{kz}$ hides the selection of a specific boundary condition for the solution of the Laplace equation (3.417). Which is it? How could different boundary conditions be set?

Inserting this result together with (3.416) into Bernoulli's equation (3.415) now immediately gives the dispersion relation for gravity waves,

$$\omega^2 = kg \ . \tag{3.418}$$

Interestingly, the group velocity $v_\mathrm{g} = \partial_k \omega$ of such waves depends on the wavelength,

$$v_\mathrm{g} = \partial_k \sqrt{kg} = \frac{1}{2}\sqrt{\frac{g}{k}} \ : \tag{3.419}$$

Longer gravity waves travel faster.

3.6.2 The Rayleigh-Taylor instability

We now consider a somewhat more involved situation. Imagine two different fluids meeting at a common, unperturbed surface perpendicular to the local gravitational acceleration g. The fluid above has density ρ_1 and height h_1, the fluid below has density ρ_2 and depth h_2. Both fluids are initially at rest. We choose the coordinate system such that the unperturbed surface coincides with the x-y plane. If this is perturbed as described by a function $\zeta(x, y, t)$, how do the perturbations develop?

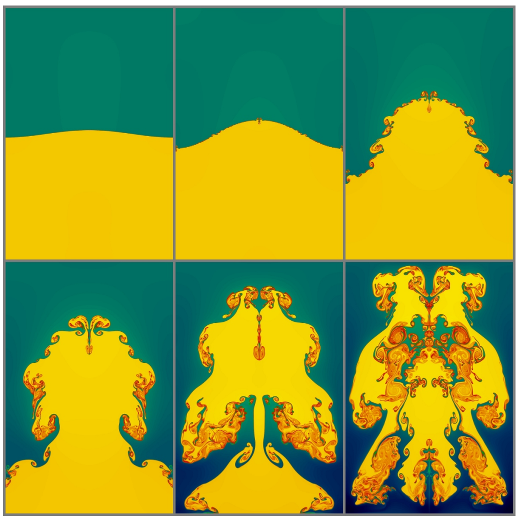

Figure 3.14 This sequence of images shows the onset of the Rayleigh-Taylor instability in a simulation. The boundaries of the rising "Rayleigh-Taylor finger" additionally become Kelvin-Helmholtz unstable. (courtesy of Volker Springel)

Now, we have to find two velocity potentials, ψ_1 and ψ_2, subject to the following conditions: Both have to satisfy the Laplace equation (3.406), the Bernoulli equation (3.415) at the surface and the boundary conditions that the pressure and the velocity at the surface must be continuous and that the velocities at $z = h_1$ and $z = -h_2$ must both vanish,

$$(P_1 - P_2)|_{z=\zeta} = 0 \,, \quad (\partial_z\psi_1 - \partial_z\psi_2)|_{z=\zeta} = 0 \,,$$
$$(\partial_z\psi_1)|_{z=h_1} = 0 = (\partial_z\psi_2)|_{z=-h_2} \,. \tag{3.420}$$

Laplace's equation, together with the third of these conditions suggests the *ansatz*

$$\psi_1 = A_1 \cosh[k(z - h_1)]e^{i(kx-\omega t)} \,,$$
$$\psi_2 = A_2 \cosh[k(z + h_2)]e^{i(kx-\omega t)} \,. \tag{3.421}$$

With Bernoulli's equation (3.412), the first of the boundary conditions (3.420) requires

$$\zeta = \frac{\rho_2\partial_t\psi_2 - \rho_1\partial_t\psi_1}{g(\rho_1 - \rho_2)} \,, \tag{3.422}$$

where we can now evaluate the right-hand side at $z = 0$ rather than at $z = \zeta$, since ζ is supposed to be a small perturbation. Another time derivative of (3.422) gives

$$\partial_t \zeta = v_z = \partial_z \psi_1 = \frac{\rho_2 \partial_t^2 \psi_2 - \rho_1 \partial_t^2 \psi_1}{g(\rho_1 - \rho_2)} \ . \tag{3.423}$$

Inserting the ansatz (3.421) here, we first find

$$-A_1 k \sinh(kh_1) = \frac{\omega^2}{g(\rho_1 - \rho_2)} \left[A_2 \rho_2 \cosh(kh_2) - A_1 \rho_1 \cosh(kh_1)\right] \ , \tag{3.424}$$

and the second boundary condition (3.420) finally requires

$$-A_1 \sinh(kh_1) = A_2 \sinh(kh_2) \ . \tag{3.425}$$

We now eliminate A_2 between the latter two equations and obtain

$$\omega^2 = \frac{kg(\rho_2 - \rho_1)}{\rho_2 \coth(kh_2) + \rho_1 \coth(kh_1)} \ . \tag{3.426}$$

This dispersion relation shows the highly interesting result that the frequency becomes imaginary if the specifically lighter fluid is placed beneath the specifically heavier one. This is the Rayleigh-Taylor or buoyancy instability (Figure 3.14): In such a configuration, small perturbations of the surface between the two fluids cause the fluids to begin exchanging their stratification.

3.6.3 The Kelvin-Helmholtz instability

We now come to another hydrodynamical instability, caused by a tangential velocity perturbation at the boundary between two fluids. Again, the boundary is described by a surface $\zeta(x, y, t)$ which, as long as the surfaces remain unperturbed, coincides with the plane $z = 0$. We proceed as follows. We imagine an unperturbed situation in which the upper fluid is streaming with velocity $\vec{v} = v\hat{e}_x$ into the x direction. We anticipate that this shear flow will excite wave-like perturbations $\delta\vec{v}$ in the velocity, δP in the pressure and $\delta z = \zeta$ in the boundary, and express this anticipation by adopting

$$\begin{aligned} &\delta\vec{v} = \delta\vec{v}_0 \, \mathrm{e}^{\mathrm{i}(kx-\omega t)} \ , \quad \delta P = \delta P_0 \, f(z) \, \mathrm{e}^{\mathrm{i}(kx-\omega t)} \ , \\ &\zeta = \zeta_0 \, \mathrm{e}^{\mathrm{i}(kx-\omega t)} \ . \end{aligned} \tag{3.427}$$

The equations we need to solve are the linearised Euler equation

$$\partial_t \delta\vec{v} + v\partial_x \delta\vec{v} = -\frac{\vec{\nabla}\delta P}{\rho} \tag{3.428}$$

for an incompressible fluid, $\vec{\nabla}\cdot\delta\vec{v} = 0$. Applying the divergence to (3.428) shows that the pressure perturbation δP now needs to satisfy the Laplace equation

$$\vec{\nabla}^2 \delta P = 0 \ , \tag{3.429}$$

which immediately implies the equation $f''(z) - k^2 f(z) = 0$ for $f(z)$ or

$$f(z) = f_0 \mathrm{e}^{-kz} \tag{3.430}$$

above the surface since the solution is then confined to $z > 0$. Now, the z component of Euler's equation shows that the velocity perturbation in z direction above the surface needs to satisfy

$$\delta v_z = \frac{k\delta P_1}{\mathrm{i}\rho_1(kv - \omega)} \, . \tag{3.431}$$

?

Notice once more the choice of a boundary condition for the Laplace equation (3.429) implied by the ansatz (3.430).

At the same time, δv_z must be given by the derivatives of the surface ζ,

$$\delta v_z = \partial_t \zeta + v\partial_x \zeta = \mathrm{i}(kv - \omega)\zeta \, . \tag{3.432}$$

Equating both expressions for δv_z, we find the relation

$$\delta P_1 = -\frac{\rho_1 \zeta}{k}(kv - \omega)^2 \tag{3.433}$$

for the pressure fluctuation above the boundary. Below the boundary, $v = 0$ and the minus sign is changed to a plus sign since $f(z) = f_0 \exp(kz)$ there. Thus,

$$\delta P_2 = \frac{\rho_2 \zeta}{k}\omega^2 \, . \tag{3.434}$$

Since the pressure fluctuation needs to be continuous at the boundary, $\delta P_1 = \delta P_2$ at $z = \zeta$,

$$\rho_2 \omega^2 + \rho_1 (kv - \omega)^2 = 0 \, . \tag{3.435}$$

This quadratic equation for ω leads to the dispersion relation

$$\omega_\pm = \frac{kv}{\rho_1 + \rho_2}\left(\rho_1 \pm \mathrm{i}\sqrt{\rho_1 \rho_2}\right) \, . \tag{3.436}$$

Unless $\rho_1 = 0$, this frequency is complex and thus necessarily implies an instability, the so-called *Kelvin-Helmholtz* instability (Figures 3.15 and 3.15).

Figure 3.15 Examples for the Kelvin-Helmholtz instability in the atmospheres of the Earth (left panel) and Jupiter (right panel), in both cases indicated by clouds. (Wikipedia)

3.6.4 Thermal Instability

Let us now consider a physical system that gains energy by heating processes and loses energy by cooling mechanisms. Both heating and cooling can occur in many ways. Frequent heating mechanisms are heating by compression, by the injection of hot particles or by radiation from nearby sources. Cooling may occur by expansion or through radiation losses, but also for example by the emission of energetic particles. The net effect of the heating and cooling processes taken together is described by the cooling function

$$\mathcal{L}(\rho, T) \, , \tag{3.437}$$

which describes the total energy loss per unit mass and unit time. It is typically a function of the density ρ and the temperature T, but may of course depend on other parameters, such as the chemical composition of the system under consideration.

If the system is in thermal equilibrium, we require that the net cooling function vanish, $\mathcal{L}(\rho, T) = 0$: thermal equilibrium requires that the rates of energy gain by heating and loss by cooling exactly balance each other. This condition implicitly defines a relation between ρ and T.

Example: Thermal bremsstrahlung

For thermal bremsstrahlung, for example, the cooling function is proportional to the squared density times the square root of the temperature,

$$\mathcal{L}(\rho, T) = C \rho \sqrt{T} - (\text{heating terms}) \, . \tag{3.438}$$

The density ρ appears linear here rather than quadratic because we refer the cooling function to unit mass rather than unit volume. More realistic cooling functions contain terms caused by so-called line cooling, i.e. cooling by the emission of energy via spectral lines. ◂

The cooling function $\mathcal{L}(\rho, T)$ can adopt various forms, in particular because cooling processes are often related to thermal occupation numbers of quantum states and the quantum-mechanical transition probabilities between atomic or molecular excitations. Since the Boltzmann factor decreases exponentially with the energy of states scaled by the thermal energy $k_\mathrm{B}T$, sometimes small temperature changes can give rise to large changes in occupation numbers. If quantum transitions contribute to the cooling processes, the discrete atomic or molecular energy levels involved introduce discrete thresholds. A typical curve in the ρ-T plane characterised by the equilibrium condition $\mathcal{L}(\rho, T) = 0$ may thus contain flat plateaus and steep steps.

As is common in thermodynamics, there may be several equilibria for the system to attain, such as mechanical, thermal or phase equilibrium. Mechanical equilibrium could be established by the system adapting to some external pressure P. In such a case, the pressure at the system's boundary is set externally. Then, the equation of state, i.e. the relation between pressure, temperature and density $P = P(\rho, T)$, may define an additional curve in the ρ-T plane. For an ideal gas, for example, we must satisfy

$$P = \frac{\rho k_\mathrm{B} T}{m_\mathrm{particle}} \, . \tag{3.439}$$

Maintaining mechanical equilibrium with a constant external pressure P_ext then defines a hyperbola in the ρ-T plane, or a straight line in the $\log\rho$-$\log T$ plane.

If thermal and mechanical equilibrium need to be maintained at the same time, the system may occupy only such points in the ρ-T plane where the curves defined by the conditions $\mathcal{L}(\rho, T) = 0$ and $P(\rho, T) = P_{\text{ext}}$ intersect. If the condition $\mathcal{L}(\rho, T) = 0$ contains plateaus and steps, this may occur in several points (Figure 3.16).

Mechanical equilibrium is usually established much faster than thermal or phase equilibria. Thus, if the external, mechanical conditions change, the system will first rearrange itself to maintain mechanical equilibrium. Doing so, it moves to another point in the ρ-T plane on the appropriate curve defined by $P(\rho, T) = P_{\text{ext}}$. This will bring it out of thermal equilibrium.

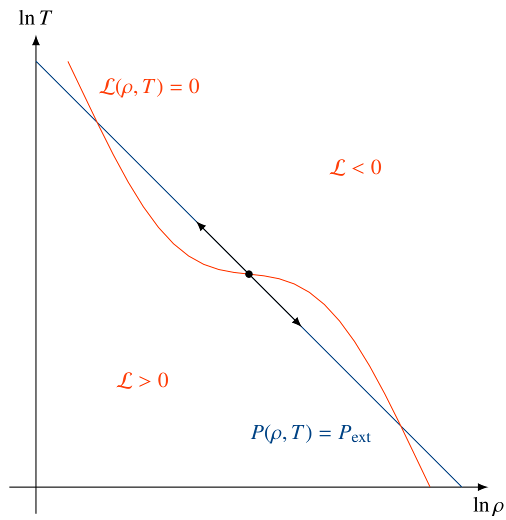

Figure 3.16 Illustration of the discussion on the onset of thermal instability in mechanical equilibrium with constant external pressure.

Example: Unstable or stable evolution

To construct a specific example, suppose that the system is heated by extra energy at constant external pressure. It will typically reestablish mechanical equilibrium faster than thermal equilibrium. To sufficient approximation, it will thus first expand and move along its isobaric curve towards lower density and higher temperature. Suppose this drives the system to a place where the cooling function is negative, $\mathcal{L} < 0$. Now, the energy gain will be larger than the energy loss. Then, the temperature will increase further, the density will decrease by further expansion, and the system will move even further away from thermal equilibrium. It will then be thermally unstable. Suppose, on the contrary, that the heating drives the system to a point where the cooling function is positive, $\mathcal{L} > 0$. It can then cool from its new position and move back to thermal equilibrium. ◂

Let us now consider a simple model for the thermal instability. Besides the continuity and Euler equations, taken in the form

$$\partial_t \rho + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \quad \text{and} \quad \partial_t \vec{v} + \frac{\vec{\nabla} v^2}{2} - \vec{v} \times \left(\vec{\nabla} \times \vec{v} \right) = -\frac{\vec{\nabla} P}{\rho} , \tag{3.440}$$

we write the energy-conservation equation

$$T\left[\partial_t s + \left(\vec{v}\cdot\vec{\nabla}\right)s\right] = -\mathcal{L}(\rho, T) \tag{3.441}$$

in terms of the specific entropy. The cooling function $\mathcal{L}$ gives the net energy loss per unit mass.

Let now the thermal equilibrium state be characterised by the density $\rho = \rho_0$ and the temperature $T = T_0$. By definition, the cooling function vanishes there, $\mathcal{L}(\rho_0, T_0) = 0$. We shall further assume that the unperturbed fluid velocity $\vec{v}_0 = 0$, which can be achieved by transforming to the comoving frame of the unperturbed flow. We perturb this state by small deviations $\delta\rho$, δT and $\delta\vec{v}$ to the density, the temperature and the velocity, respectively, and linearise in these perturbations. The linearised continuity and Euler equations read

$$\partial_t\delta\rho + \vec{\nabla}\cdot(\rho_0\delta\vec{v}) = 0\,, \quad \partial_t\delta\vec{v} = -\frac{\vec{\nabla}\delta P}{\rho_0}\,. \tag{3.442}$$

As usual, we eliminate the divergence of the velocity perturbation by combining the time derivative of the continuity equation with the divergence of the Euler equation. This enables us to write

$$\partial_t^2\delta\rho = \vec{\nabla}^2\delta P\,. \tag{3.443}$$

We first allow perturbations with $\delta P \neq 0$ and ask later for the conditions for instability under constant pressure.

We continue by linearising the entropy equation. We expand the specific entropy near its equilibrium value s_0 as

$$s = s_0 + \frac{\partial s}{\partial P}\delta P + \frac{\partial s}{\partial\rho}\delta\rho = s_0 + c_\mathrm{v}\frac{\delta P}{P_0} - c_\mathrm{p}\frac{\delta\rho}{\rho_0}\,, \tag{3.444}$$

where the earlier result (3.250) was used in the second step. Likewise, we expand the cooling function on the right-hand side of (3.441),

$$\mathcal{L}(\rho, T) = \mathcal{L}_0 + \left(\frac{\partial\mathcal{L}}{\partial T}\right)_\rho \delta T_\rho + \left(\frac{\partial\mathcal{L}}{\partial T}\right)_P \delta T_P\,, \tag{3.445}$$

distinguishing temperature changes at constant density, δT_ρ, and at constant pressure, δT_P. For later convenience, we introduce the abbreviations

$$\mathcal{L}_P := \frac{1}{c_\mathrm{p}}\left(\frac{\partial\mathcal{L}}{\partial T}\right)_P \quad \text{and} \quad \mathcal{L}_V := \frac{1}{c_\mathrm{v}}\left(\frac{\partial\mathcal{L}}{\partial T}\right)_\rho \tag{3.446}$$

for the derivatives of the cooling function with respect to temperature, taken at constant pressure or constant density. Since the specific entropy s_0 at equilibrium must satisfy the unperturbed entropy equation (3.441), we can insert the expansions (3.444) and (3.445) into (3.441), eliminate the equilibrium expressions and linearise in the perturbations to obtain

$$T_0\partial_t\left(c_\mathrm{v}\frac{\delta P}{P_0} - c_\mathrm{p}\frac{\delta\rho}{\rho_0}\right) = -c_\mathrm{v}\mathcal{L}_V\delta T_\rho - c_\mathrm{p}\mathcal{L}_P\delta T_P\,. \tag{3.447}$$

Now, for an ideal gas, the perturbed equation of state is

$$\delta P = \frac{\rho_0 k_B T_0}{\bar{m}} \left(\frac{\delta T}{T_0} + \frac{\delta\rho}{\rho_0} \right) = P_0 \left(\frac{\delta T}{T_0} + \frac{\delta\rho}{\rho_0} \right) , \tag{3.448}$$

implying that the temperature changes at constant density and at constant pressure are

$$\delta T_\rho = T_0 \frac{\delta P}{P_0} \quad \text{and} \quad \delta T_P = -T_0 \frac{\delta\rho}{\rho_0} . \tag{3.449}$$

These expressions allow us to bring (3.447) into the form

$$\partial_t \left(c_\mathrm{v} \frac{\delta P}{P_0} - c_\mathrm{p} \frac{\delta\rho}{\rho_0} \right) = c_\mathrm{p} \mathcal{L}_P \frac{\delta\rho}{\rho_0} - c_\mathrm{v} \mathcal{L}_V \frac{\delta P}{P_0} . \tag{3.450}$$

After multiplying with P_0, replacing

$$\frac{P_0}{\rho_0} = \frac{c_\mathrm{s}^2}{\gamma} \tag{3.451}$$

according to (3.240) and recalling $c_\mathrm{p} = \gamma c_\mathrm{v}$, we arrive at

$$\partial_t \left(\delta P - c_\mathrm{s}^2 \delta\rho \right) = c_\mathrm{s}^2 \mathcal{L}_P \delta\rho - \mathcal{L}_V \delta P . \tag{3.452}$$

At this point, we take the Laplacian and insert the result (3.443) obtained previously from the linearly perturbed continuity and Euler equations. This leads to the equation

$$\partial_t \left(\partial_t^2 \delta\rho - c_\mathrm{s}^2 \vec{\nabla}^2 \delta\rho \right) = c_\mathrm{s}^2 \mathcal{L}_P \vec{\nabla}^2 \delta\rho - \mathcal{L}_V \partial_t^2 \delta\rho \tag{3.453}$$

for the density perturbations $\delta\rho$.

As usual in linear stability analysis, we evaluate this equation for a plane wave

$$\delta\rho = \delta\hat{\rho} \mathrm{e}^{\mathrm{i}(kx - \omega t)} , \tag{3.454}$$

for which (3.453) turns into

$$\partial_t \left[\left(c_\mathrm{s}^2 k^2 - \omega^2 \right) \delta\rho \right] = \left(\mathcal{L}_V \omega^2 - \mathcal{L}_P c_\mathrm{s}^2 k^2 \right) \delta\rho , \tag{3.455}$$

which yields the cubic dispersion relation

$$-\mathrm{i}\omega \left(c_\mathrm{s}^2 k^2 - \omega^2 \right) = \mathcal{L}_V \omega^2 - \mathcal{L}_P c_\mathrm{s}^2 k^2 . \tag{3.456}$$

In general, this equation is difficult to solve. In the limiting case of small wave lengths, $c_\mathrm{s}^2 k^2 \gg \omega^2$, we can approximate

$$\mathrm{i}\omega \approx \mathcal{L}_P \quad \text{or} \quad \omega \approx -\mathrm{i}\mathcal{L}_P . \tag{3.457}$$

Then, the density perturbation depends on time as

$$\delta\rho = \delta\hat{\rho}\, \mathrm{e}^{-\mathcal{L}_P t} , \tag{3.458}$$

hence it grows exponentially if $\mathcal{L}_P < 0$. Thermal instability thus sets in on small scales if

$$\mathcal{L}_P = \frac{1}{c_\mathrm{p}} \left(\frac{\partial \mathcal{L}}{\partial T} \right)_P < 0 , \tag{3.459}$$

i.e. if the cooling function decreases upon temperature increases at constant pressure. In the opposite limiting case of very large wave length, $c_s^2 k^2 \ll \omega^2$, the dispersion relation (3.456) demands that

$$\omega \approx -\mathrm{i}\mathcal{L}_V \ . \tag{3.460}$$

Thermal intability then sets in if

$$\mathcal{L}_V = \frac{1}{c_\mathrm{v}} \left(\frac{\partial \mathcal{L}}{\partial T} \right)_\rho < 0 \ . \tag{3.461}$$

The conditions (3.459) and (3.461) show that thermal instability must be expected if the cooling function decreases upon temperature increases, or in other words, if higher temperature leads to reduced cooling or conversely if lower temperature implies enhanced cooling. These conditions are of course quite intuitive: If a system can cool more efficiently the cooler it gets, or if it can cool less efficiently the hotter it gets, any small temperature fluctuation can be expected to grow.

3.6.5 Heat conduction

Since mechanical equilibrium can typically be established much faster than thermal equilibrium, systems can be in mechanical equilibrium, but out of thermal equilibrium. Perhaps the most straightforward example is a star which is being kept in mechanical equilibrium by the balance between gravity and the pressure gradient, but nonetheless continuously radiates energy. In this case, a temperature gradient is maintained between the core and the surface by the central energy production, and the entropy equation reads

$$\rho T \frac{\mathrm{d}s}{\mathrm{d}t} = \vec{\nabla} \cdot \left(\kappa \vec{\nabla} T \right) + \sigma_{ij} \frac{\partial v^i}{\partial x_j} \ . \tag{3.462}$$

If the velocity gradient on the right-hand side can be considered too small to drive any matter currents, the second term on the right-hand side can be neglected. At constant pressure, we can write the change δq of heat per unit mass as

$$\delta q = c_\mathrm{p} \mathrm{d}T = T \mathrm{d}s \ , \tag{3.463}$$

which we can solve for the differential

$$\mathrm{d}s = c_\mathrm{p} \mathrm{d} \ln T \tag{3.464}$$

of the specific entropy. This allows us to reduce the entropy equation (3.462) to read

$$\rho c_\mathrm{p} \frac{\mathrm{d}T}{\mathrm{d}t} = \kappa \vec{\nabla}^2 T \ . \tag{3.465}$$

Introducing the transport coefficient $\chi \equiv \kappa/(\rho c_\mathrm{p})$ for the temperature, we can re-write the entropy equation as a diffusion equation

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \chi \vec{\nabla}^2 T \tag{3.466}$$

for the temperature T with the diffusion coefficient χ.

In close analogy to radiative energy transport, in particular the result (2.484) for the radiative energy-current density, we now define a conductive opacity $\kappa_{\rm cond}$ through the conductive energy current

$$\vec{F}_{\rm cond} = -\frac{c}{3\rho\kappa_{\rm cond}}\vec{\nabla}\left(aT^4\right) = -\frac{4acT^3}{3\rho\kappa_{\rm cond}}\vec{\nabla}T \equiv -\kappa\vec{\nabla}T \,, \tag{3.467}$$

from which we obtain the relation

$$\kappa = \frac{4caT^3}{3\rho\kappa_{\rm cond}} \tag{3.468}$$

between the heat conductivity κ and the conductive opacity $\kappa_{\rm cond}$. If both radiative and conductive energy transport are present, an effective opacity $\kappa_{\rm eff}$ can thus usefully be defined by

$$\frac{1}{\kappa_{\rm eff}} = \frac{1}{\kappa_{\rm rad}} + \frac{1}{\kappa_{\rm cond}} \quad \Rightarrow \quad \kappa_{\rm eff} = \frac{\kappa_{\rm rad}\kappa_{\rm cond}}{\kappa_{\rm rad} + \kappa_{\rm cond}} \,. \tag{3.469}$$

?

Why is it useful to define the inverse of the effective opacity as the sum of the inverse radiative and conductive opacities as in (3.469)?

3.6.6 Convection

We have just seen that temperature gradients drive energy currents of either by electromagnetic radiation or heat conduction. If the temperature gradient is too large, convection sets in. Then, warm, rising bubbles cannot cool sufficiently and adapt to their environment as they rise. Instead, they remain warmer than the surrounding medium and continue to rise. We now investigate this situation, considering a volume $V(P, s)$ of gas characterised by the pressure P and the specific entropy s as it rises against the gravitational force (Figure 3.17).

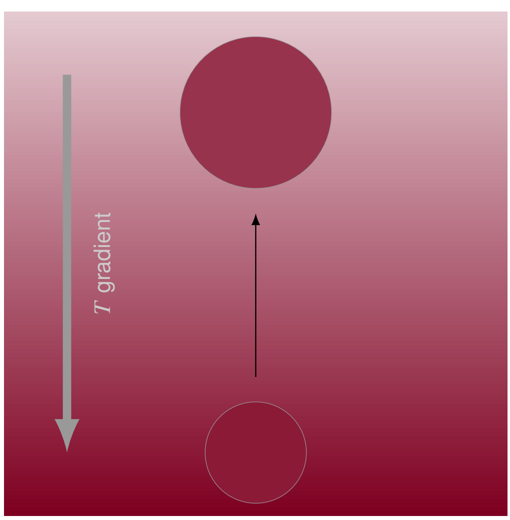

Figure 3.17 Convection sets in if rising bubbles retain higher entropy than their environment.

We ignore again thermal compared to mechanical adaptation processes because they are typically slower. We watch the bubble with volume $V(P, s)$ as it rises by

Example: Heat conductivity due to electrons

When we discussed the heat conductivity as a phenomenon due to particle transport, we saw in (3.135) that it can be written as

$$\kappa = \frac{nv}{\sqrt{3}} c_v \lambda \tag{3.470}$$

in terms of the mean-free path λ of the fluid particles. If we consider electrons whose mean free path is determined by scattering off the ions,

$$\lambda = \frac{1}{n_i \sigma} , \tag{3.471}$$

where n_i and σ are the number density of the ions and their scattering cross section with the electrons.

Typically, an electron will approach an ion up to a distance r_i where the kinetic and potential energies equal,

$$\frac{mv^2}{2} \approx \frac{Ze^2}{r_i} \quad \Rightarrow \quad r_i \approx \frac{2Ze^2}{mv^2} . \tag{3.472}$$

The cross section for electron-ion scattering can then be crudely approximated by

$$\sigma \approx \pi r_i^2 , \tag{3.473}$$

and we obtain the expression

$$\kappa = \frac{n_e v_e}{\sqrt{3}} c_v \frac{m_e^2 v_e^4}{4\pi n_i Z^2 e^4} = \frac{1}{\sqrt{3}} \left(\frac{m_e^2}{4\pi Z^2 e^4} \right) \left(\frac{n_e}{n_i} \right) c_v v_e^5 \tag{3.474}$$

for the heat conductivity contributed by electrons scattered by ions.

In a thermal electron gas, the heat capacity at constant volume is $c_v = 3k_B/2$ per particle and the thermal electron velocity is

$$v_e^2 = \frac{3k_B T_e}{m_e} . \tag{3.475}$$

Inserted into (3.474), these results give the heat conductivity

$$\kappa = \frac{\sqrt{3} k_B}{2} \left(\frac{m_e^2}{4\pi Z^2 e^4} \right) \left(\frac{n_e}{n_i} \right) \left(\frac{3k_B T_e}{m_e} \right)^{5/2} \tag{3.476}$$

for classical (non-degenerate) electrons. ◀

Example: Heat conductivity due to electrons (continued)

If we identify the Thomson cross section σ_T here, we can alternatively write

$$\kappa = \frac{9k_B c}{Z^2\sigma_T}\left(\frac{n_e}{n_i}\right)\left(\frac{k_B T_e}{m_e}\right)^{5/2} , \tag{3.477}$$

which obviously has the required dimension

$$[\kappa] = \frac{\mathrm{erg}}{\mathrm{cm\,s\,K}} . \tag{3.478}$$

Numerically, we find

$$\kappa \approx 9.5 \cdot 10^{12} Z^{-2}\left(\frac{n_e}{n_i}\right)\left(\frac{kT_e}{1\,\mathrm{keV}}\right)^{5/2} . \tag{3.479}$$

◂

an amount Δz, where its volume after the essentially instantaneous mechanical adaptation is $V' = V(P', s)$. Having risen, the bubble experiences a buoyancy force determined by the volume $V'' = V(P', s')$ which the bubble *would* adopt if it had the specific entropy s' of its new environment. This situation is stable if the actual bubble volume $V' = V(P', s)$ is smaller than the adapted volume $V'' = V(P', s')$, because then gravity will dominate the buoyancy force, and the bubble will then sink down again. We thus have the condition

$$V(P', s') = V'' > V' = V(P', s) \tag{3.480}$$

for convective stability.

The entropy at the increased height $z + \Delta z$ is

$$s' = s + \left.\frac{\mathrm{d}s}{\mathrm{d}z}\right|_z \Delta z , \tag{3.481}$$

and the volume change of the bubble with specific entropy at constant pressure is

$$\mathrm{d}V = \left(\frac{\partial V}{\partial s}\right)_P \mathrm{d}s = c_p \left(\frac{\partial V}{\partial s}\right)_P \frac{\mathrm{d}T}{T} . \tag{3.482}$$

In its new environment at increased height $z + \Delta z$ with its specific entropy s', the bubble thus attains the new volume

$$V'' = V' + \left(\frac{\partial V}{\partial s}\right)_P \Delta s = V' + \left(\frac{\partial V}{\partial s}\right)_P \left.\frac{\mathrm{d}s}{\mathrm{d}z}\right|_z \Delta z . \tag{3.483}$$

The stability condition (3.480) is thus satisfied if the specific entropy increases with the height z,

$$\left.\frac{\mathrm{d}s}{\mathrm{d}z}\right|_z > 0 . \tag{3.484}$$

The derivative of the specific entropy with respect to the height z can be expressed by

$$\frac{\mathrm{d}s}{\mathrm{d}z} = \left(\frac{\partial s}{\partial T}\right)_P \frac{\mathrm{d}T}{\mathrm{d}z} + \left(\frac{\partial s}{\partial P}\right)_T \frac{\mathrm{d}P}{\mathrm{d}z} = c_v\left[\gamma\frac{\mathrm{d}\ln T}{\mathrm{d}z} - (\gamma - 1)\frac{\mathrm{d}\ln P}{\mathrm{d}z}\right] , \tag{3.485}$$

where we have used the partial derivatives (3.246) of the entropy found earlier. The stability condition (3.484) then shows that the temperature gradient must satisfy

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln z} > \frac{\gamma - 1}{\gamma}\frac{\mathrm{d}\ln P}{\mathrm{d}\ln z} \tag{3.486}$$

for the gas stratification to be *unstable* against convection. The quantity

$$\frac{\gamma - 1}{\gamma} \equiv \nabla_{\mathrm{ad}} \tag{3.487}$$

is often called the adiabatic temperature gradient ("nabla adiabatic"). Using this, we stability condition is written in the compact form

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} \equiv \nabla < \nabla_{\mathrm{ad}} \ . \tag{3.488}$$

?

Why is ∇_{ad} from (3.487) called "adiabatic" temperature gradient? To see this, work out

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}$$

for an adiabatically stratified gas.

Once convection sets in, it is a very efficient means of transporting heat, but viscosity can hinder the convective energy transport.

3.6.7 Turbulence

Turbulence is a very rich, important and fascinating field of its own. By no means can we treat turbulence here in any depth. We confine the discussion here to one physically and methodically aspect, i.e. the derivation of the Kolmogorov power spectrum for the scale dependence of energy in subsonic turbulence.

Hydrodynamical flows with large Reynolds numbers turn out to be highly unstable. For high viscosity (low Reynolds number), stable solutions of the Navier-Stokes equation exist which develop instabilities above a critical Reynolds number

$$\mathcal{R} = \frac{uL}{\nu} \gtrsim \mathcal{R}_{\mathrm{cr}} \ . \tag{3.489}$$

A full analysis of such instabilities is very difficult and in general an unsoved problem. Turbulence sets in, in the course of which energy is being transported from large to small scales until it is dissipated by the production of viscous heat on sufficiently small scales. Turbulence can be seen as the transitional regime between macroscopic, ordered as opposed to microscopic, unordered or thermal motion. On the macroscopic scale, turbulence is driven by some stirring mechanism acting on some linear scale. Eddies form on that scale which feed a cascade of eddies of decreasing size. This proceeds until the smallest eddies reach a size comparable to the mean-free path of the fluid particles. The energy fed into the turbulent cascade on the driving scale propagates through the cascade of eddies and is finally dissipated into heat by dissipation, i.e. by the viscosity of the fluid.

Let λ be the size of an eddy within the turbulent cascade of the fluid flow, and v_λ the linear velocity that the eddy rotates with. The characteristic time scale of one turn-over of the eddy is $\tau_\lambda = \lambda/v_\lambda$. Let further ε_λ be the specific energy, i.e.

the energy per unit mass characteristic for the material in such an eddy. Then, the flow of the specific energy through an eddy of the size λ is

$$\dot{\varepsilon} \approx \frac{\mathrm{d}\varepsilon}{\mathrm{d}\tau} \approx \overbrace{\left(\frac{v_\lambda^2}{2}\right)}^{\text{specific energy}} \underbrace{\left(\frac{\lambda}{v_\lambda}\right)^{-1}}_{\text{inverse time scale}} \approx \frac{v_\lambda^3}{\lambda} \,. \tag{3.490}$$

Let L be the macroscopic length scale on which the energy is being fed into the turbulent cascade, and u be the typical stirring velocity. From there, the energy cascades through the turbulent eddies to progressively smaller scales until it is finally viscously dissipated on a scale λ_{visc}. In between, i.e. on scales λ satisfying the scale hierarchy

$$\lambda_{\mathrm{visc}} < \lambda < L \,, \tag{3.491}$$

the energy flow $\dot{\varepsilon}$ must be independent of scale because the energy cannot be accumulated at any scale in the cascade. Therefore, we conclude from (3.490) that the typical eddy velocity must change with the eddy scale λ as

$$v_\lambda \propto \lambda^{1/3} \,. \tag{3.492}$$

Together with the boundary condition that the velocity be u on the driving scale L, we thus expect

$$v_\lambda \approx u\left(\frac{\lambda}{L}\right)^{1/3} \,. \tag{3.493}$$

The largest eddies thus rotate with the highest velocities, but the smallest have the highest vorticity,

$$\Omega \approx \frac{v_\lambda}{\lambda} \approx \frac{u}{\lambda}\left(\frac{\lambda}{L}\right)^{1/3} \approx \frac{u}{(\lambda^2 L)^{1/3}} \,. \tag{3.494}$$

To estimate the viscous scale λ_{visc}, we compare the viscous dissipation with the specific energy flow $\dot{\varepsilon}$. The viscous heating rate h_{visc} can be estimated by the viscosity times the squared velocity gradient, as can be seen from the dissipation term on the right-hand side of the Navier-Stokes equation (3.148). Thus,

$$h_{\mathrm{visc}} \approx \eta\left(\frac{v_\lambda}{\lambda}\right)^2 \approx \eta\left(\frac{v_\lambda^3}{\lambda}\right)^{2/3} \lambda^{-4/3} = \eta\dot{\varepsilon}^{2/3}\lambda^{-4/3} \,. \tag{3.495}$$

Therefore, h_{visc} is negligibly small on large scales, but if the heating rate becomes of the order of the energy flow rate,

$$h_{\mathrm{visc}} \approx \rho\dot{\varepsilon} \,, \tag{3.496}$$

viscous dissipation begins dominating. According to (3.495), this happens on a length scale λ_{visc} given by

$$\eta\dot{\varepsilon}^{2/3}\lambda_{\mathrm{visc}}^{-4/3} \approx \rho\dot{\varepsilon} \,. \tag{3.497}$$

Solving for λ_{visc} and inserting $\dot{\varepsilon} = u^3/L$ gives

$$\lambda_{\mathrm{visc}} = \left(\frac{\eta L^{1/3}}{\rho u}\right)^{3/4} = L\left(\frac{\nu}{uL}\right)^{3/4} = \frac{L}{\mathcal{R}^{3/4}} \,, \tag{3.498}$$

where $\mathcal{R}$ is the Reynolds number on the scale L.

Finally, we consider how the specific energy is distributed over scales. Doing so, we assess how the specific energy scales with the wave number k. Since the squared velocity scales like $\lambda^{2/3}$, its Fourier transform scales like

$$\widehat{\left(v_\lambda^2\right)} \propto \lambda^3\,\lambda^{2/3} \propto \lambda^{11/3} \propto k^{-11/3}\;. \tag{3.499}$$

The number of Fourier modes in shells of width dk is $\propto k^2$, hence the energy spectrum as a function of wave number is

$$E(k) \propto k^2\,k^{-11/3} = k^{-5/3}\;. \tag{3.500}$$

This is the famous energy spectrum derived by Kolmogorov in 1941, showing how the energy in a turbulent cascade is distributed over scales identified by their wave number k.

Problems

1. Studying gravity waves, we have solved (3.417) with the implicitly assumed boundary condition $f(z) = 0$ for $z \to -\infty$. The dispersion relation (3.418) is thus valid for infinitely deep water.
   (a) Derive the solution of (3.417) satisfying the boundary condition $f(z) = 0$ at finite depth, $z = -h$.
   (b) Which dispersion relation do you find for water with finite depth?

Suggested further reading: [10, 11, 12, 13, 14, 15, 16, 17]

Chapter 4

Fundamentals of Plasma Physics and Magnetohydrodynamics

4.1 Collision-less Plasmas

We begin this chapter on plasma physics and magnetohydrodynamics with a discussion of Debye shielding, showing that charges embedded in a plasma have a Yukawa- rather than a Coulomb potential with a characteristic length scale, the Debye length (4.11), which can be combined with the mean thermal velocity of the plasma particles to derive the plasma frequency (4.16).

4.1.1 Shielding and the Debye length

Plasmas are gases whose particles are charged. They typically occur when the kinetic energies of the gas particles exceed the ionisation energy of the atomic species they are composed of. The atoms are then partially or fully ionised. Unless the positive and the negative charges, i.e. the ions and the electrons, are separated by macroscopic electric fields, the plasma is macroscopically neutral. In subvolumes of a plasma whose linear dimensions are much larger than the typical inter-particle separation, there are thus on average as many negative as positive charges.

For many purposes, a plasma can be treated as a single fluid. This is possible if not only the mean-free path for collisions between the plasma particles is much smaller than any macroscopically relevant length scale, but if also the interactions between the positive and the negative charges are fast enough for them not to separate on a macroscopic scale. Sometimes, however, the positive and the negative charges need to be treated as two interacting fluids.

In principle, the electromagnetic interaction between the positive and the negative charges has an infinite range because the Coulomb force falls off like the inverse-squared distance from a charge. The Coulomb interaction between its particles is thus the fundamental difference between plasma physics and the hydrodynamics of neutral fluids. A treatment of plasmas as fluids, however, requires that collisions between the plasma particles be random, short-ranged and fast, such that equilibrium can locally be quickly established. This is possible despite the Coulomb interaction because the existence of two different types of charge allows shielding on a characteristic length scale which we first want to work out.

Let the plasma consist of electrons of charge $-e$ and ions of charge Ze. The spatial number densities of the electrons and the ions be n_e and n_i, respectively. We begin with a macroscopically neutral plasma with negative charge density $-en_\mathrm{e}$ and positive charge density Zen_i. For the plasma to be neutral, the number densities must be related by

$$Zen_\mathrm{i} = en_\mathrm{e} \quad \Rightarrow \quad n_\mathrm{i} = \frac{n_\mathrm{e}}{Z} \; . \tag{4.1}$$

Suppose now we place a point charge q at the (arbitrary) origin into the plasma. This point charge will displace the positive and negative plasma charges to some degree and thereby change their number densities locally. If q is positive, as we shall assume without loss of generality, n_i will be lowered in its immediate neighbourhood, while n_e will be slightly increased there compared to the equilibrium densities of the electrons and the ions given by (4.1). The local imbalance between the positive and the negative charge distributions will create an electrostatic potential Φ different from zero.

The thermal motion of the plasma particles will counteract their displacement. In presence of an electrostatic potential, the particles will rearrange such as to minimise their potential energies $Ze\Phi$ and $-e\Phi$. The unperturbed equilibrium densities $\bar{n}_\mathrm{e}$ for the negative charges and $\bar{n}_\mathrm{i} = \bar{n}_\mathrm{e}/Z$ for the positive charges will thus be modified by a Boltzmann factor and read

$$n_\mathrm{i} = \frac{\bar{n}_\mathrm{e}}{Z} \exp\left(\frac{Ze\Phi}{k_\mathrm{B}T}\right) \; , \quad n_\mathrm{e} = \bar{n}_\mathrm{e} \exp\left(-\frac{e\Phi}{k_\mathrm{B}T}\right) \; . \tag{4.2}$$

The potential Φ is determined by the Poisson equation

$$\begin{aligned} \vec{\nabla}^2\Phi &= 4\pi\left(Zen_\mathrm{i} - en_\mathrm{e}\right) + 4\pi q\delta_\mathrm{D}\left(\vec{x}\right) \\ &= 4\pi\bar{n}_\mathrm{e}e\left[\exp\left(\frac{Ze\Phi}{k_\mathrm{B}T}\right) - \exp\left(\frac{-e\Phi}{k_\mathrm{B}T}\right)\right] + 4\pi q\delta_\mathrm{D}\left(\vec{x}\right) \end{aligned} \tag{4.3}$$

together with the boundary condition that $\Phi \to 0$ far away from the point charge. On the right-hand side of the Poisson equation, the point charge q appears at the coordinate origin in addition to the plasma charges. If q is not too large and the plasma is not very cold, it is appropriate to assume

$$\frac{e\Phi}{k_\mathrm{B}T} \ll 1 \; , \quad \frac{Ze\Phi}{k_\mathrm{B}T} \ll 1 \tag{4.4}$$

and to Taylor-expand the exponentials in (4.3) to first order. This leads us to the approximate Poisson equation

$$\vec{\nabla}^2\Phi = 4\pi(Z+1)\,\frac{\bar{n}_\mathrm{e}e^2}{k_\mathrm{B}T}\,\Phi + 4\pi q\delta_\mathrm{D}\left(\vec{x}\right) \; , \tag{4.5}$$

which is most easily solved after transforming it into Fourier space. Introducing the *Debye wave number*

$$k_{\rm D}^2 = 4\pi\frac{\bar{n}_{\rm e}e^2}{k_{\rm B}T}\;, \tag{4.6}$$

we can write the Fourier-transformed Poisson equation as

$$\hat{\Phi}(k) = -\frac{4\pi q}{(Z+1)k_{\rm D}^2 + k^2}\;. \tag{4.7}$$

Caution The approximate Poisson equation (4.5) is an inhomogeneous Helmholtz equation. ◀

This is straightforwardly transformed back into real space because $\hat{\Phi}(k)$ depends on the wave number only, but not on the direction of the wave vector. The angular integrations in the inverse Fourier transform first give

$$\begin{aligned}\Phi(r) &= 2\pi\int_0^\infty \frac{k^2{\rm d}k}{(2\pi)^3}\,\hat{\Phi}(k)\int_{-1}^1 {\rm d}\cos\theta\,{\rm e}^{{\rm i}kr\cos\theta} \\ &= 4\pi\int_0^\infty \frac{k^2{\rm d}k}{(2\pi)^3}\,\hat{\Phi}(k)\,\frac{\sin(kr)}{kr}\;. \end{aligned}\tag{4.8}$$

With the help of the definite integral

$$\int_0^\infty {\rm d}x\,\frac{x\sin(\alpha x)}{\beta^2+x^2} = \frac{\pi}{2}{\rm e}^{-\alpha\beta}\;, \tag{4.9}$$

the remaining k integration in (4.9) over the potential (4.7) can now directly be carried out to give

$$\Phi(r) = -\frac{q}{r}\exp\left(-\sqrt{Z+1}k_{\rm D}r\right)\;. \tag{4.10}$$

? How could you prove (4.9)?

By the presence of the plama charges, the Coulomb potential of the point charge q is thus changed to a Yukawa potential (Figure 4.1) which decreases exponentially on the typical length scale

$$\lambda_{\rm D} = k_{\rm D}^{-1} = \left(\frac{k_{\rm B}T}{4\pi\bar{n}_{\rm e}e^2}\right)^{1/2}\;. \tag{4.11}$$

Notice that $\lambda_{\rm D}$ does not depend on the charge q placed into the plasma! This is the *Debye length*, which gives the characteristic length scale for the shielding of an arbitrary charge in an otherwise neutral plasma (Figure 4.2a).

A plasma is called *ideal* if it contains many particles in a volume given by the cubed Debye length. Then the interaction energy between positive and negative charges is small compared to their thermal energy; in other words, the electrostatic interactions affect the thermal motion of the plasma particles only very weakly. To see this, we compare the mean potential energy $Ze^2/\bar{r}$ of an electron in the Coulomb field of an ion with its kinetic energy $3k_{\rm B}T/2$ in thermal equilibrium.

The mean separation $\bar{r}$ between the particles is determined by

$$\frac{4\pi}{3}\bar{r}^3\bar{n}_{\rm e} \approx 1 \quad\Rightarrow\quad \bar{r} \approx \left(\frac{3}{4\pi\bar{n}_{\rm e}}\right)^{1/3}\;. \tag{4.12}$$

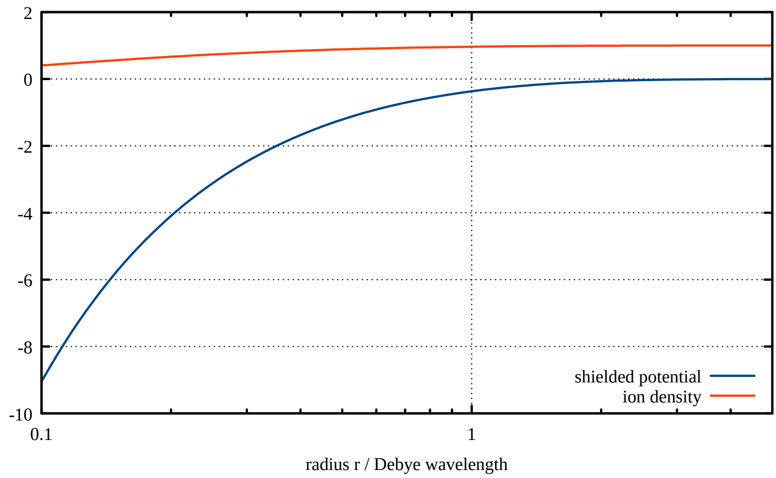

Figure 4.1 A point charge in a plasma rearranges the surrounding charges. Its electrostatic potential is shielded and thus exponentially cut off.

The ratio between the electrostatic interaction energy and the thermal energy is thus (Figure 4.2b)

$$\frac{Ze^2}{k_\mathrm{B}T}\left(\frac{4\pi\bar{n}_\mathrm{e}}{3}\right)^{1/3} = Z\frac{4\pi\bar{n}_\mathrm{e}e^2}{3k_\mathrm{B}T}\left(\frac{3}{4\pi\bar{n}_\mathrm{e}}\right)^{2/3} = \frac{Z}{3}\frac{\bar{r}^2}{\lambda_\mathrm{D}^2}\,, \tag{4.13}$$

which is much less than unity if the Debye length greatly exceeds the mean inter-particle separation $\bar{r}$. For an ideal plasma, we can thus assume $\bar{r} \ll \lambda_\mathrm{D}$ by definition.

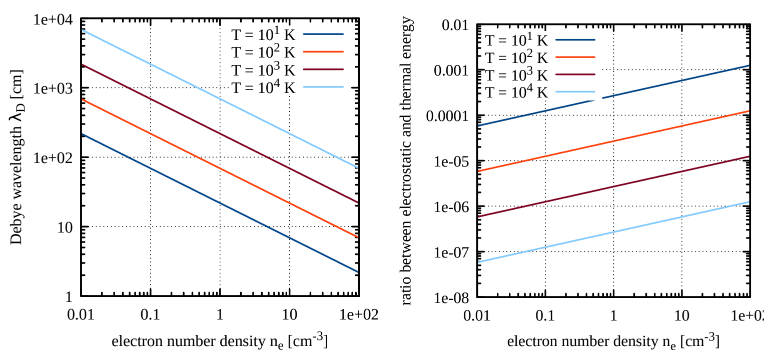

Figure 4.2 The Debye wavelength λ_D in centimetres (left panel) and the ratio between the electrostatic and the thermal energies (right panel) are shown as functions of the plasma temperature T and the mean electron number density $\bar{n}_\mathrm{e}$.

4.1.2 The plasma frequency

By the equipartition theorem from statistical physics, the mean-squared thermal velocity of the plasma electrons in one spatial direction is

$$\left\langle v^2 \right\rangle = \frac{k_\mathrm{B} T}{m_\mathrm{e}} \tag{4.14}$$

if they are in thermal equilibrium with a plasma at temperature T. A plasma electron thus moves by the Debye length in a mean time interval of

$$t_\mathrm{D} = \frac{\lambda_\mathrm{D}}{\langle v^2 \rangle^{1/2}} = \sqrt{\frac{k_\mathrm{B} T}{4\pi \bar{n}_\mathrm{e} e^2} \frac{m_\mathrm{e}}{k_\mathrm{B} T}} = \sqrt{\frac{m_\mathrm{e}}{4\pi \bar{n}_\mathrm{e} e^2}} \,. \tag{4.15}$$

This sets the time scale on which the thermal motion of the electrons can compensate charge displacements by shielding. This characteristic reaction time t_D of the plasma can be transformed into a characteristic frequency,

$$\omega_\mathrm{p} = \frac{1}{t_\mathrm{D}} = \sqrt{\frac{4\pi \bar{n}_\mathrm{e} e^2}{m_\mathrm{e}}} \approx 5.6 \cdot 10^4\,\mathrm{Hz} \left(\frac{\bar{n}_\mathrm{e}}{\mathrm{cm}^{-3}} \right)^{1/2} \,, \tag{4.16}$$

called the *plasma frequency*. External changes applied to the plasma with frequencies higher than the plasma frequency, for example by an incident electromagnetic wave, are too fast for the plasma particles to adapt and rearrange, while changes with lower frequency can be accommodated. With the Debye length λ_D and the plasma frequency ω_p, we now have two essential parameters at hand for describing plasmas.

Problems

1. Solve the Poisson equation (4.5) directly, i.e. without transforming into Fourier space. *Hint*: Solve the homogeneous equation first, introducing spherical polar coordinates. Then solve the inhomogeneous equation by variation of constants. For solving the homogeneous equation, try the ansatz

$$\Phi(r) = \Phi_0 (\alpha r)^n \exp(-\alpha r) \tag{4.17}$$

and see whether a suitable exponent n can be found.

4.2 Electromagnetic Waves in Media

This section is a recollection of electrodynamics in media, introducing the dielectric displacement $\vec{D}$, the polarisation $\vec{P}$ and the dielectric tensor ε. For later convenience, we decompose in (4.37) the dielectric tensor into its components parallel and perpendicular to the propagation direction of electromagnetic waves and define the longitudinal and transverse dielectricities in (4.40).

4.2.1 Polarisation and dielectric displacement

An important part of plasma physics concerns the propagation of electromagnetic waves through plasmas. There, Maxwell's vacuum equations in vacuum no longer hold because the plasma as a medium reacts to the presence of electromagnetic fields, and thereby alters them. While the homogeneous Maxwell equations remain unchanged, the inhomogeneous equations change due to the appearance of charges and currents that only appear because the external electromagnetic fields act on the charges and possible magnetic dipoles of the medium.

Recall that electric and magnetic fields are defined by forces on test charges and test-current loops. Such test systems are idealisations whose own, intrinsic fields are so small that the fields can be considered unchanged that are supposed to be measured. The electric force experienced by a test charge embedded in a medium is expressed by the dielectric displacement $\vec{D}$ instead of the electric field $\vec{E}$, while a test-current loop in the medium experiences a magnetic force expressed by the magnetic field strength $\vec{H}$ instead of the magnetic field $\vec{B}$.

Considering the linearity of Maxwell's equations, it is natural to assume that $\vec{D}$ and $\vec{H}$ be linearly related to $\vec{E}$ and $\vec{B}$, respectively. We thus adopt the common linear relations

$$\vec{D} = \varepsilon\vec{E} \ , \quad \vec{B} = \mu\vec{H} \ , \tag{4.18}$$

where the dielectricity ε and the magnetic permeability μ appear. In the Gaussian system of units, $\vec{D}$ and $\vec{H}$ are defined such that ε and μ of the vacuum are unity. Even though the relations (4.18) appear very simple, several complications may arise. First, ε and μ may depend on space and time. Second, they may be tensors if the medium has preferred spatial directions imprinted, such as the principal axes of a crystal or the magnetic field lines in a magnetised plasma. Then, not only the magnitudes of $\vec{D}$ and $\vec{H}$ may differ from those of $\vec{E}$ and $\vec{B}$, but also their directions if the principal axes of the ε and μ tensors are misaligned with $\vec{E}$ and $\vec{B}$.

?

Can you construct an (artificial) dielectric tensor ε such that $\vec{D}$ is perpendicular to $\vec{E}$?

In what follows, we shall generally consider astrophysical media whose particles have no relevant magnetic moments. Then, external magnetic fields will neither be dimished nor strengthened by macroscopically aligned, microscopic magnetic dipoles, the medium will not respond to the presence of an external magnetic field, and we can identify $\vec{B}$ with $\vec{H}$, adopting $\mu = 1$. We shall thus only consider the response of the medium to external electric fields in the following.

External electric fields $\vec{E}$ polarise media, i.e. they slightly displace the positive and negative charges of the microscopic constituents of these media. To lowest, but sufficient order in a multipole expansion, these charge displacements can be described by electric dipole moments. When macroscopically averaged over scales large compared to the individual particles, but small compared to the overall dimensions of the complete medium, these microscopic dipoles can be linearly superposed to form the macroscopic polarisation $\vec{P}$, which counteracts the external electric field $\vec{E}$. The divergence of the polarisation corresponds to a polarised charge density $\rho_{\rm pol}$, defined by

$$\rho_{\rm pol} = -\vec{\nabla} \cdot \vec{P} \ . \tag{4.19}$$

Inside the medium, this polarised charged density must be added to any free charge density ρ that may additionally be present. The Maxwell equation for an electric field in vacuum, $\vec{\nabla} \cdot \vec{E} = 4\pi\rho$, must now be augmented by the polarised charge density $\rho_{\rm pol}$,

$$\vec{\nabla} \cdot \vec{E} = 4\pi \left(\rho + \rho_{\rm pol}\right) = 4\pi\rho - 4\pi\vec{\nabla} \cdot \vec{P} \ . \tag{4.20}$$

As discussed before, the dielectric displacement, $\vec{D} \equiv \vec{E} + 4\pi\vec{P}$, is introduced as an auxiliary field describing the response of the medium to an external electric field. Its sources are the free charges only,

$$\vec{\nabla} \cdot \vec{D} = 4\pi\rho \ , \tag{4.21}$$

excluding any charges $\rho_{\rm pol}$ that are exclusively caused by polarisation of otherwise neutral microscopic particles.

Charge conservation, expressed by the continuity equation for the charge density, must also apply to the polarised charge density,

$$\frac{\partial\rho_{\rm pol}}{\partial t} + \vec{\nabla} \cdot \vec{j}_{\rm pol} = 0 \ . \tag{4.22}$$

If we substitute (4.19) here, we find the current density $j_{\rm pol}$ caused by changes of the polarisation with time,

$$\vec{\nabla} \cdot \left[-\frac{\partial\vec{P}}{\partial t} + \vec{j}_{\rm pol}\right] = 0 \quad \text{or} \quad \vec{j}_{\rm pol} = \frac{\partial\vec{P}}{\partial t} \ , \tag{4.23}$$

where the final step assumes that $j_{\rm pol}$ is curl-free. This current density needs to be added to the current density $\vec{j}$ of the free charges. The induction equation then reads

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\frac{\partial\vec{E}}{\partial t} + \frac{4\pi}{c}\left(\vec{j} + \vec{j}_{\rm pol}\right) = \frac{1}{c}\left(\frac{\partial\vec{E}}{\partial t} + 4\pi\frac{\partial\vec{P}}{\partial t}\right) + \frac{4\pi}{c}\vec{j} \tag{4.24}$$

and can be brought into the form

$$\vec{\nabla} \times \vec{B} = \frac{1}{c}\partial_t\vec{D} + \frac{4\pi}{c}\vec{j} \ : \tag{4.25}$$

In a medium, it is not the time derivative of the electric field, but of the dielectric displacement that creates the curl of the magnetic field together with the current. Equations (4.21) and (4.25) replace the previous inhomogeneous Maxwell equations for the divergence of the electric field $\vec{E}$ and the curl of the magnetic field $\vec{B}$.

In the following, we shall assume macroscopically neutral media in which the free charge density vanishes, $\rho = 0$. This does not necessarily imply that there could be no macroscopic currents. In fact, the free current density $\vec{j}$ may differ from zero. We shall further assume that the microscopic constituents of the media have no net magnetic moment, allowing us to neglect any distinction between $\vec{B}$ and $\vec{H}$. In such media, Maxwell's equations then read

$$\vec{\nabla} \cdot \vec{D} = 0 \ , \quad \dot{\vec{D}} + 4\pi\vec{j} = c\vec{\nabla} \times \vec{H} \ , \quad \dot{\vec{B}} = -c\vec{\nabla} \times \vec{E} \ , \quad \vec{\nabla} \cdot \vec{B} = 0 \ , \tag{4.26}$$

augmented by the linear relations

$$\vec{D} = \varepsilon\vec{E} \;, \quad \vec{j} = \sigma\vec{E} \tag{4.27}$$

between $\vec{D}$, $\vec{E}$ and $\vec{j}$, where the dielectricity ε and the conductivity σ may be tensors. For the time being, we shall neglect free currents, setting $\vec{j} = 0$. Later on, in magnetohydrodynamics, we shall have to take them into account.

4.2.2 Structure of the dielectric tensor

The dielectric displacement $\vec{D}$ may be delayed with respect to the external field $\vec{E}$, and it may propagate through the medium. The dielectric displacement $\vec{D}(t, \vec{x})$ at a time t and a place $\vec{x}$ may thus depend on the electric field $\vec{E}(t', \vec{x}')$ at an earlier time $t' < t$ and another place $\vec{x}'$. The dielectric displacement would then be determined by a convolution of the electric field with a suitable kernel because they must still be related in a linear way. Since convolutions are multiplications in Fourier space, it is thus reasonable to begin with a multiplicative ansatz in Fourier space.

Second, the medium itself may have a preferred direction. This could be a crystal axis or the local direction of a magnetic field. Then, $\vec{D}$ does no longer need to be colinear with $\vec{E}$, but may point into a different direction. Such a change of orientation can be expressed by introducing a dielectric tensor instead of a dielectric constant.

Taking these two arguments together, we begin with

$$\hat{\vec{D}}\left(\omega, \vec{k}\right) = \hat{\varepsilon}\left(\omega, \vec{k}\right) \hat{\vec{E}}\left(\omega, \vec{k}\right) \;, \tag{4.28}$$

defining the *dielectric tensor* $\hat{\varepsilon}(\omega, \vec{k})$ as a function of frequency ω and wave vector $\vec{k}$. Since the fields must remain real in configuration space $(t, \vec{x})$, the dielectric tensor must satisfy the symmetry relation

$$\hat{\varepsilon}\left(-\omega, -\vec{k}\right) = \hat{\varepsilon}^*\left(\omega, \vec{k}\right) \tag{4.29}$$

in Fourier space $(\omega, \vec{k})$.

If the medium, in our case the plasma, does not imprint a specific direction, the only vector that can be used to span the tensor $\hat{\varepsilon}^i_{\;j}$ is the wave vector $\vec{k}$ of an incoming electromagnetic wave itself. We introduce a unit vector $\hat{k}$ in the direction of $\vec{k}$ by $\hat{k} = \vec{k}/k$ and begin with the ansatz

$$\hat{\varepsilon}\left(\omega, \vec{k}\right) = \hat{A}\mathbb{1}_3 + \hat{B}\,\hat{k} \otimes \hat{k} \tag{4.30}$$

with functions $\hat{A}(\omega, k)$ and $\hat{B}(\omega, k)$ that may depend on the frequency ω and the wave number k. That is, we linearly compose the dielectric tensor $\hat{\varepsilon}$ of the unit tensor and the tensor $\hat{k} \otimes \hat{k}$, the only basis tensors we have available here (Figure 4.3). Notice that the tensor $\hat{\varepsilon}$ defined this way is symmetric. In particular, this means that $\vec{D}$ and $\vec{E}$ can be multiplied in a symmetric scalar product,

$$\hat{\vec{D}} \cdot \hat{\vec{E}} = \hat{\vec{E}}^\top \hat{\varepsilon} \hat{\vec{E}} \;. \tag{4.31}$$

If we had reasons to believe that $\hat{\varepsilon}$ should contain an antisymmetric part, this could be supplied by adding a contribution proportional to the Levi-Civita tensor, e.g. of the form

$$\varepsilon_{ijk}\hat{k}^k \ . \tag{4.32}$$

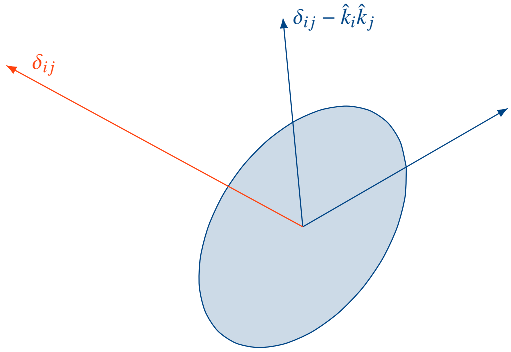

Figure 4.3 The direction of the wave vector $\vec{k}$ itself defines the structure of the dielectric tensor, if no other preferred directions are defined in the plasma.

Obviously, since $\vec{k}\cdot\hat{k} = \vec{k}^2/k = k$, the expression

$$\left(\hat{k}\otimes\hat{k}\right)\vec{k} =: \pi_{\|}\vec{k} = \vec{k} \tag{4.33}$$

is parallel to $\vec{k}$, while

$$\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k}\right)\vec{k} =: \pi_{\perp}\vec{k} = \vec{k} - \vec{k} = 0 \tag{4.34}$$

vanishes and is thus perpendicular to $\vec{k}$. The tensors $\pi_{\|}$ and $\pi_{\perp}$ with the components

$$\pi_{\|} = \hat{k}\otimes\hat{k} \quad \text{and} \quad \pi_{\perp} = \mathbb{1}_3 - \hat{k}\otimes\hat{k} = \mathbb{1}_3 - \pi_{\|} \tag{4.35}$$

are generally convenient *projectors* for vector components parallel and perpendicular to the direction $\hat{k}$ satisfying

$$\pi_{\|}^2 = \pi_{\|} \ , \quad \pi_{\perp}^2 = \pi_{\perp} \ , \quad \pi_{\|}\pi_{\perp} = 0 = \pi_{\perp}\pi_{\|} \ . \tag{4.36}$$

?

As remarked before, projections need to be idempotent, in the present case $\pi_{\|}^2 = \pi_{\|}$ and $\pi_{\perp}^2 = \pi_{\perp}$. Confirm and interpret equations (4.36) geometrically.

We can use them we split the tensor $\hat{\varepsilon}$ into a transversal and a longitudinal part,

$$\hat{\varepsilon} = \hat{\varepsilon}_{\perp}\pi_{\perp} + \hat{\varepsilon}_{\|}\pi_{\|} \ , \tag{4.37}$$

where the transversal and the longitudinal dielectricities $\hat{\varepsilon}_{\perp}$ and $\hat{\varepsilon}_{\|}$ were defined. These are related to the functions $\hat{A}$ and $\hat{B}$ introduced in (4.30) above by $\hat{A} = \hat{\varepsilon}_{\perp}$ and $\hat{B} = \hat{\varepsilon}_{\|} - \hat{\varepsilon}_{\perp}$. Of course, $\hat{\varepsilon}_{\perp}$ and $\hat{\varepsilon}_{\|}$ are generally functions of ω and k which also need to satisfy the symmetry condition (4.29),

$$\hat{\varepsilon}_{\perp,\|}(-\omega, k) = \hat{\varepsilon}^*_{\perp,\|}(\omega, k) \ . \tag{4.38}$$

Contracting (4.37) with either of the projection tensors $\pi^{\perp}$ or $\pi^{\|}$, and using that their traces are

$$\mathrm{Tr}\,\pi_{\perp} = 2 \quad \text{and} \quad \mathrm{Tr}\,\pi_{\|} = 1 \ , \tag{4.39}$$

we can project the longitudinal and transverse dielectricities out of the dielectric tensor,

$$\hat{\varepsilon}_\perp = \frac{1}{2}\,\mathrm{Tr}\,\pi_\perp\hat{\varepsilon} \quad \text{and} \quad \hat{\varepsilon}_\parallel = \mathrm{Tr}\,\pi_\parallel\hat{\varepsilon} \ . \tag{4.40}$$

Notice explicitly that we have neglected in this decomposition of the dielectric tensor $\hat{\varepsilon}_{ij}(\omega, \vec{k})$ that preferred macroscopic directions may exist in the plasma, e.g. due to magnetic fields ordered on large scales. If they exist, they must also be built into the dielectric tensor.

?

Why is there a factor of 1/2 in the expression for $\hat{\varepsilon}_\perp$ in (4.40)?

Problems

1. In media, the electric and magnetic fields $\vec{E}$ and $\vec{B}$ in Maxwell's inhomogeneous equations need to be replaced according to (4.18), while Maxwell's homogeneous equations remain unchanged.

   (a) From Maxwell's equations in media and Ohm's law $\vec{j} = \sigma\vec{E}$, derive the telegraph equation

   $$\vec{\nabla}^2\vec{E} - \frac{\varepsilon\mu}{c^2}\partial_t^2\vec{E} = \frac{4\pi\sigma\mu}{c^2}\partial_t\vec{E} \ . \tag{4.41}$$

   Assume plane-wave solutions for $\vec{E}$ and derive the dispersion relation for waves in media.

   (b) From the equation of motion

   $$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \frac{e}{m}\vec{E} - \frac{\vec{v}}{\tau} \tag{4.42}$$

   for the electrons, containing a damping term with a characteristic collision time τ, derive an equation for the current density $\vec{j}$. Assume harmonic time dependence of $\vec{E}$ and $\vec{j}$ and identify the conductivity

   $$\sigma = \frac{ne^2}{m}\frac{\tau}{1 - \mathrm{i}\omega\tau} \ . \tag{4.43}$$

   (c) Combine the results from the preceding subproblems, assume $\mu = 1$ and $\omega\tau \gg 1$ and identify the plasma frequency ω_p. What does the limit $\omega\tau \gg 1$ mean?

4.3 Dispersion Relations

We proceed in this section by deriving the general expression (4.51) for the dispersion relation of electromagnetic waves in a plasma, which we split into the two dispersion relations (4.53) and (4.54) for transverse and longitudinal waves. By a perturbative analysis of the one-particle phase-space distribution of the plasma charges and its evolution equation, we derive the models (4.69) and (4.69) for the longitudinal and the transverse dielectricity. We conclude by deriving the Landau damping rate (4.87) of longitudinal waves.

4.3.1 General form of the dispersion relations

The dielectric tensor determines which kinds of electromagnetic wave can propagate through the plasma. The conditions for propagating waves are given by dispersion relations, which relate the frequency ω to the wave vector $\vec{k}$. Recall that the dispersion relation for electromagnetic waves in vacuum is $c^2\omega^2 = \vec{k}^2$. Based on our ansatz for the dielectric tensor and its decomposition into a longitudinal and a transversal part, we shall now derive the dispersion relations for electromagnetic waves propagating through a plasma.

We begin as usual by decomposing the incoming waves into plane waves with a phase factor $\exp[\mathrm{i}(\vec{k}\cdot\vec{x} - \omega t)]$. When applied to plane waves, Maxwell's equations (4.26) in a medium read

$$\vec{k}\times\hat{\vec{E}} = \frac{\omega\hat{\vec{B}}}{c}\;, \quad \vec{k}\times\hat{\vec{B}} = -\frac{\omega\hat{\vec{D}}}{c}\;,$$
$$\vec{k}\cdot\hat{\vec{D}} = 0\;, \quad \vec{k}\cdot\hat{\vec{B}} = 0\;, \tag{4.44}$$

if we neglect any free charge densities and currents. Combining the curl of the first equation (4.44) with the second yields

$$\vec{k}\times\left(\vec{k}\times\hat{\vec{E}}\right) = \frac{\omega}{c}\,\vec{k}\times\hat{\vec{B}} = -\frac{\omega^2}{c^2}\hat{\vec{D}}\;. \tag{4.45}$$

If we expand the double vector product, we see that the dielectric displacement vector must satisfy the equation

$$\frac{\omega^2}{c^2}\hat{\vec{D}} = k^2\hat{\vec{E}} - \vec{k}\left(\vec{k}\cdot\hat{\vec{E}}\right)\;. \tag{4.46}$$

We now introduce the dielectric tensor by substituting $\hat{\vec{D}} = \hat{\varepsilon}\hat{\vec{E}}$ and find

$$\frac{\omega^2}{c^2}\hat{\varepsilon}\hat{\vec{E}} = k^2\hat{\vec{E}} - \left(\vec{k}\otimes\vec{k}\right)\hat{\vec{E}} = k^2\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k}\right)\hat{\vec{E}} \tag{4.47}$$

or, after dividing by k^2 and rearranging,

$$\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k} - \frac{\omega^2}{c^2k^2}\,\hat{\varepsilon}\right)\hat{\vec{E}} = \left(\pi_\perp - \frac{\omega^2}{c^2k^2}\,\hat{\varepsilon}\right)\hat{\vec{E}} = 0\;. \tag{4.48}$$

This is a linear equation with the expression in parentheses representing a square, 3×3 matrix. Equation (4.48) has non-trivial solutions $\hat{\vec{E}} \neq 0$ if and only if the determinant of this matrix vanishes,

$$\det\left(\pi_\perp - \frac{\omega^2}{c^2k^2}\,\hat{\varepsilon}\right) = 0\;. \tag{4.49}$$

This is the general form of the dispersion relation between the frequency ω and the wave vector $\vec{k}$ for electromagnetic waves that can propagate through the plasma.

It is now convenient to insert the decomposition (4.37) of the dielectric tensor and to rewrite the matrix in (4.48) such that its transverse and longitudinal components are grouped together. This results in

$$\left[\left(1 - \frac{\omega^2}{c^2k^2}\hat{\varepsilon}_\perp\right)\pi_\perp - \frac{\omega^2}{c^2k^2}\hat{\varepsilon}_\parallel\pi_\parallel\right]\hat{\vec{E}} = 0 \tag{4.50}$$

and the form

$$\det\left[\left(1-\frac{\omega^2}{c^2k^2}\hat{\varepsilon}_\perp\right)\pi_\perp-\frac{\omega^2}{c^2k^2}\hat{\varepsilon}_\|\pi_\|\right]=0 \tag{4.51}$$

of the dispersion relation.

4.3.2 Transversal and longitudinal waves

We expect that the last condition (4.51) defines more than one dispersion relation because the dielectric tensor $\hat{\varepsilon}$ has a longitudinal and a transversal part. In contrast to electromagnetic waves in vacuum, which are exclusively transversal, longitudinal as well as transversal electromagnetic waves may occur in media.

For transversal waves, the projection of $\hat{\vec{E}}$ on $\vec{k}$ vanishes, $\pi_\|\hat{\vec{E}} = 0$, while $\pi_\perp\hat{\vec{E}} = \hat{\vec{E}}$. The matrix equation (4.50) then reduces to the simpler equation

$$\left(1-\frac{\omega^2}{c^2k^2}\,\hat{\varepsilon}_\perp\right)\hat{\vec{E}}=0\ , \tag{4.52}$$

which implies the dispersion relation

$$\omega^2=\frac{c^2k^2}{\hat{\varepsilon}_\perp}\ . \tag{4.53}$$

Caution Note that, according to the dispersion relation (4.53), the refractive index

$$n_\perp=\hat{\varepsilon}_\perp^{1/2}$$

for transversal electromagnetic waves can be assigned to a magnetised plasma. ◀

This recovers the usual result that transversal electromagnetic waves propagate in a medium with a reduced velocity $c\hat{\varepsilon}_\perp^{-1/2}$.

For longitudinal waves, $\pi_\perp\hat{\vec{E}} = 0$ and $\pi_\|\hat{\vec{E}} = \hat{\vec{E}}$, and the matrix equation (4.50) is reduced to

$$\frac{\omega^2}{c^2k^2}\hat{\varepsilon}_\|\hat{\vec{E}}=0\ . \tag{4.54}$$

Generally, this requires that the longitudinal dielectricity itself must vanish, $\hat{\varepsilon}_\| = 0$. In order to understand this condition, we first need to determine the form of the longitudinal and transversal dielectricities, $\hat{\varepsilon}_\|$ and $\hat{\varepsilon}_\perp$.

4.3.3 Longitudinal and transversal dielectricities

In order to determine the form of $\hat{\varepsilon}_\|$ and $\hat{\varepsilon}_\perp$, we invoke the collision-less Boltzmann equation to study the response of the plasma particles to the incoming electromagnetic wave. We neglect the motion of the ions because of their lower velocities and concentrate on the plasma electrons. Before the electromagnetic wave arrives, the phase space density is assumed to have attained an equilibrium value f_0 which is then slightly perturbed by the wave,

$$f=f_0+\delta f\ . \tag{4.55}$$

This expresses our expectation that sufficiently weak fields $\vec{E}$ and $\vec{B}$ will perturb the phase-space distribution function only by a small amount away from the equilibrium distribution f_0. Inserting the perturbation ansatz (4.55) into the

Boltzmann equation, subtracting the pure equilibrium terms and dropping terms of second order in the perturbation then gives

$$\frac{\partial \delta f}{\partial t} + \vec{v} \cdot \vec{\nabla} \delta f - e\left(\vec{E} + \frac{\vec{v}}{c} \times \vec{B}\right) \cdot \frac{\partial f_0}{\partial \vec{p}} = 0 \,, \tag{4.56}$$

where $\vec{v}$ and $\vec{p}$ are the equilibrium plasma velocity and momentum. For a locally isotropic distribution f_0, we must further have

$$\frac{\partial f_0}{\partial \vec{p}} \parallel \vec{v} \tag{4.57}$$

because no other preferred directions can be present. Thus

$$\left(\vec{v} \times \vec{B}\right) \cdot \frac{\partial f_0}{\partial \vec{p}} = 0 \,, \tag{4.58}$$

and Boltzmann's equation in linear approximation shrinks to

$$\frac{\partial \delta f}{\partial t} + \vec{v} \cdot \vec{\nabla} \delta f = e\vec{E} \cdot \frac{\partial f_0}{\partial \vec{p}} \,. \tag{4.59}$$

We now decompose the incoming electric field into plane waves with a phase factor $\exp[\mathrm{i}(\vec{k} \cdot \vec{x} - \omega t)]$ and assume that the phase-space distribution will respond in form of plane waves with equal phase. Then, (4.59) turns into the algebraic equation

$$-\mathrm{i}\omega \delta f + \mathrm{i}\vec{v} \cdot \vec{k} \delta f = e\vec{E} \cdot \frac{\partial f_0}{\partial \vec{p}} \,, \tag{4.60}$$

which can be solved for the perturbation δf of the phase-space distribution,

$$\delta f = -\frac{\mathrm{i}e\vec{E}}{\vec{k} \cdot \vec{v} - \omega} \cdot \frac{\partial f_0}{\partial \vec{p}} \,. \tag{4.61}$$

If the equilibrium distribution f_0 is locally homogeneous, isotropic and stationary, charge and current densities are exclusively caused by the perturbations δf of f_0. Therefore, the polarised charge density ρ_{pol} and the polarised current density $\vec{j}_{\mathrm{pol}}$ are

$$\rho_{\mathrm{pol}} = -e \int \mathrm{d}^3 p \, \delta f \,, \quad \vec{j}_{\mathrm{pol}} = -e \int \mathrm{d}^3 p \, \delta f \vec{v} \,. \tag{4.62}$$

These quantities are then also proportional to the same phase factor $\exp[\mathrm{i}(\vec{k} \cdot \vec{x} - \omega t)]$, and the polarisation equations (4.19) and (4.23) can be written in the form

$$\mathrm{i}\vec{k} \cdot \hat{\vec{P}} = -\hat{\rho}_{\mathrm{pol}} \,, \quad -\mathrm{i}\omega \hat{\vec{P}} = \hat{\vec{j}}_{\mathrm{pol}} \,. \tag{4.63}$$

We take the second of these equations and insert δf from (4.61) to find first

$$-\mathrm{i}\omega \hat{\vec{P}} = \mathrm{i}e^2 \int \mathrm{d}^3 p \, \frac{\vec{v}}{\vec{k} \cdot \vec{v} - \omega} \frac{\partial f_0}{\partial \vec{p}} \cdot \hat{\vec{E}} \,. \tag{4.64}$$

Since at the same time we have to satisfy the general relation

$$4\pi \hat{\vec{P}} = \hat{\vec{D}} - \hat{\vec{E}} = (\hat{\varepsilon} - \mathbb{1}_3)\, \hat{\vec{E}} \,, \tag{4.65}$$

we can directly read the dielectric tensor $\hat{\varepsilon}$ off (4.64),

$$\hat{\varepsilon} = \mathbb{1}_3 - \frac{4\pi e^2}{\omega} \int \mathrm{d}^3 p \, \frac{\vec{v}}{\vec{k} \cdot \vec{v} - \omega} \otimes \frac{\partial f_0}{\partial \vec{p}} \; . \tag{4.66}$$

By means of the projection tensors, we can now project out the longitudinal and transversal components of the dielectric tensor, as shown in (4.40). In this way, we first find the longitudinal dielectricity

$$\hat{\varepsilon}_{\parallel} = \mathrm{Tr}\,\pi_{\parallel}\hat{\varepsilon} = 1 - \frac{4\pi e^2}{\omega k^2} \int \mathrm{d}^3 p \, \frac{\vec{k} \cdot \vec{v}}{\vec{k} \cdot \vec{v} - \omega} \frac{\partial f_0}{\partial \vec{p}} \cdot \vec{k} \; . \tag{4.67}$$

Noticing that

$$\frac{\vec{k} \cdot \vec{v}}{\vec{k} \cdot \vec{v} - \omega} = 1 + \frac{\omega}{\vec{k} \cdot \vec{v} - \omega} \; , \tag{4.68}$$

and integrating once by parts, we can bring this expression into the form

$$\hat{\varepsilon}_{\parallel} = 1 - \frac{4\pi e^2}{k^2} \int \mathrm{d}^3 p \, \frac{1}{\vec{k} \cdot \vec{v} - \omega} \frac{\partial f_0}{\partial \vec{p}} \cdot \vec{k} \; . \tag{4.69}$$

The transverse dielectricity is

$$\hat{\varepsilon}_{\perp} = \frac{1}{2}\,\mathrm{Tr}\,\pi_{\perp}\hat{\varepsilon} = 1 - \frac{2\pi e^2}{\omega} \int \mathrm{d}^3 p \, \frac{1}{\vec{k} \cdot \vec{v} - \omega} \frac{\partial f_0}{\partial \vec{p}_{\perp}} \cdot \vec{v}_{\perp} \; , \tag{4.70}$$

where the perpendicular velocity $\vec{v}_{\perp}$ and the perpendicular momentum $\vec{p}_{\perp}$ are defined by

$$\vec{v}_{\perp} = \pi_{\perp}\vec{v} = \vec{v} - \left(\hat{k} \cdot \vec{v}\right)\hat{k} \; , \quad \vec{p}_{\perp} = m\vec{v}_{\perp} \; . \tag{4.71}$$

?

Beginning with (4.66), confirm the expressions (4.69) and (4.69) for the transverse and longitudinal dielectric tensors.

4.3.4 Landau Damping

Before we proceed to calculate the longitudinal and the transversal dielectricities for the special, but frequent case of a thermal plasma, we consider longitudinal waves in particular and identify an interesting damping process.

Because of the pole at $\vec{k} \cdot \vec{v} = \omega$, the longitudinal dielectricity $\hat{\varepsilon}_{\parallel}$ has a real and an imaginary part. The latter is responsible for damping of the incoming waves because it leads to an imaginary frequency. As we shall see, this damping process dissipates the incoming electromagnetic energy. To begin with, we note that the energy dissipation Q rate has two contributions, one from the damping of the electromagnetic waves and the associated decrease of the electromagnetic field energy density, and another from the Ohmic heating,

$$Q = \frac{\partial}{\partial t}\left(\frac{\vec{E}^2}{8\pi}\right) + \vec{E} \cdot \vec{j}_{\mathrm{pol}} \; . \tag{4.72}$$

In absence of any current of the free charges, the current density is solely the polarisation current defined by the continuity equation (4.22) and related to the polarisation change by (4.23). We can thus continue calculating the dissipation as

$$Q = \frac{\vec{E} \cdot \dot{\vec{E}}}{4\pi} + \vec{E} \cdot \dot{\vec{P}} = \frac{\vec{E}}{4\pi}\left(\dot{\vec{E}} + 4\pi\dot{\vec{P}}\right) = \frac{\vec{E} \cdot \dot{\vec{D}}}{4\pi} \; . \tag{4.73}$$

Notice that the electromagnetic field, the dielectric displacement and the polarisation do not wear hats here: They are to be taken as functions of $\vec{x}$ and t here.

We now consider the contribution of an individual plane wave characterised by frequency ω and wave number $\vec{k}$ to the dissipation Q, i.e. we insert

$$\vec{E} = \hat{\vec{E}}\,\mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)}\;, \quad \vec{D} = \hat{\vec{D}}\,\mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)} \tag{4.74}$$

into (4.73). The dissipation Q resulting therefrom will thus be the dissipation *per Fourier mode* $(\omega, \vec{k})$. Since we should finally arrive at a real expression for Q, we replace $\vec{E}$ by $(\vec{E} + \vec{E}^*)/2$ to obtain its real part, and likewise for $\vec{D}$ by $(\vec{D} + \vec{D}^*)/2$. Furthermore, since $\hat{\vec{D}} = \hat{\varepsilon}_{\|}\hat{\vec{E}}$ for longitudinal waves, we can write

$$\hat{\vec{D}} = \hat{\varepsilon}_{\|}\hat{\vec{E}} \quad \rightarrow \quad \frac{1}{2}\left(\hat{\varepsilon}_{\|}\hat{\vec{E}} + \hat{\varepsilon}_{\|}^*\hat{\vec{E}}^*\right)\;. \tag{4.75}$$

Inserting these expressions for $\vec{E}$ and $\vec{D}$ into Q from (4.73) gives

$$Q = -\frac{\mathrm{i}\omega}{16\pi}\left(\vec{E} + \vec{E}^*\right)\left(\hat{\varepsilon}_{\|}\vec{E} - \hat{\varepsilon}_{\|}^*\vec{E}^*\right)\;, \tag{4.76}$$

where the minus sign in the second factor comes from the change in sign in the phase factor $\exp[\mathrm{i}(\vec{k}\cdot\vec{x} - \omega t)]$ due to the complex conjugation of the dielectric displacement. Averaging the dissipation (4.76) over time eliminates the products $\vec{E}\cdot\vec{E}$ and $\vec{E}^*\cdot\vec{E}^*$ because they vary with the phase factor like $\exp(-2\mathrm{i}\omega t)$, while the mixed terms become independent of time. Thus, the time-averaged dissipation is

$$\langle Q\rangle = -\frac{\mathrm{i}\omega}{16\pi}\left(\hat{\varepsilon}_{\|}\hat{\vec{E}}\cdot\hat{\vec{E}}^* - \hat{\varepsilon}_{\|}^*\hat{\vec{E}}^*\cdot\hat{\vec{E}}\right) = -\frac{\mathrm{i}\omega}{16\pi}\left(\hat{\varepsilon}_{\|} - \hat{\varepsilon}_{\|}^*\right)\left|\hat{\vec{E}}\right|^2\;. \tag{4.77}$$

The remaining expression in brackets is twice the imaginary part of the longitudinal dielectricity $\hat{\varepsilon}_{\|}$,

$$\hat{\varepsilon}_{\|} - \hat{\varepsilon}_{\|}^* = 2\mathrm{i}\,\mathrm{Im}\,\hat{\varepsilon}_{\|}\;, \tag{4.78}$$

and thus we find

$$\langle Q\rangle = \frac{\omega}{8\pi}\,\mathrm{Im}\,\hat{\varepsilon}_{\|}\left|\hat{\vec{E}}\right|^2\;. \tag{4.79}$$

The imaginary part of $\hat{\varepsilon}_{\|}$ can be obtained from (4.71). In order to avoid the pole in the integrand there, we shift it away from the real axis by a small amount δ,

$$\frac{1}{\vec{k}\cdot\vec{v} - \omega} \rightarrow \frac{1}{\vec{k}\cdot\vec{v} - \omega - \mathrm{i}\delta}\;, \tag{4.80}$$

and then take the imaginary value

$$\mathrm{Im}\,\frac{1}{\vec{k}\cdot\vec{v} - \omega - \mathrm{i}\delta} = \mathrm{Im}\,\frac{\vec{k}\cdot\vec{v} - \omega + \mathrm{i}\delta}{\left(\vec{k}\cdot\vec{v} - \omega\right)^2 + \delta^2} = \frac{\delta}{\left(\vec{k}\cdot\vec{v} - \omega\right)^2 + \delta^2}\;. \tag{4.81}$$

In the limit $\delta \rightarrow 0$, this turns into π times a Dirac delta function,

$$\lim_{\delta\rightarrow 0}\frac{\delta}{\left(\vec{k}\cdot\vec{v} - \omega\right)^2 + \delta^2} = \pi\delta_{\mathrm{D}}\left(\vec{k}\cdot\vec{v} - \omega\right)\;. \tag{4.82}$$

This can be seen by verifying that the limit satisfies the two defining criteria of a δ function,

$$\lim_{\delta\to 0}\frac{\delta}{\pi\left(x^2+\delta^2\right)}=0\quad (x\neq 0) \tag{4.83}$$

and

$$\lim_{\delta\to 0}\int_{-R}^{R}\frac{\delta}{\pi\left(x^2+\delta^2\right)}\mathrm{d}x=\frac{2}{\pi}\lim_{\delta\to 0}\arctan\frac{R}{\delta}=1\;. \tag{4.84}$$

Without loss of generality, we now rotate the coordinate frame such that the x axis aligns with the wave vector $\vec{k}$. We further integrate the one-particle phase-space distribution function f_0 over the momentum components p_y and p_z and define

$$\bar{f}(p_x)=\int \mathrm{d}p_y\int \mathrm{d}p_z\, f_0\left(\vec{p}\right)\;. \tag{4.85}$$

The imaginary part of the longitudinal dielectricity $\hat{\varepsilon}_{\parallel}$ can then be written as

$$\begin{aligned}\mathrm{Im}\,\hat{\varepsilon}_{\parallel}&=-\frac{4\pi^2e^2}{k^2}\int \mathrm{d}p_x\,k\frac{\mathrm{d}\bar{f}}{\mathrm{d}p_x}\,\delta_{\mathrm{D}}\left(kv_x-\omega\right)\\&=-\frac{4\pi^2e^2m_{\mathrm{e}}}{k^2}\left.\frac{\mathrm{d}\bar{f}}{\mathrm{d}p_x}\right|_{p_x=\omega m/k}\;,\end{aligned} \tag{4.86}$$

and the mean dissipation rate turns into

$$\langle Q\rangle=-\left|\hat{\vec{E}}\right|^2\frac{\pi m e^2\omega}{k^2}\left.\frac{\mathrm{d}\bar{f}}{\mathrm{d}p_x}\right|_{p_x=\omega m/k}\;. \tag{4.87}$$

This is *Landau damping*, which is caused by the fact that electrons which are slightly faster than the wave are slowed down, electrons which are slightly slower than the wave are accelerated, and since the velocity distribution is typically monotonically decreasing, more electrons need to be accelerated than decelerated, and thus the wave loses energy.

Problems

1. Landau damping in a thermal plasma:

   (a) Evaluate the mean energy dissipation rate $\langle Q\rangle$ due to Landau damping for waves propagating through a thermal plasma, assuming (as will be shown in the next section) that longitudinal waves with the plasma frequency can propagate.

   (b) Estimate a time scale for Landau damping of a longitudinal wave in a thermal plasma, depending on its wavelength.

4.4 Electromagnetic Waves in Thermal Plasmas

In this section, we evaluate the microscopic models for the longitudinal and the transverse dielectricities from the preceding section for a thermal plasma. This leads us to the expressions (4.101) and (4.102), giving the two dielectricities in terms of the plasma dispersion function. From appropriate expansions of this function, approximations to the dispersion relations for frequencies low and high compared to the plasma frequency are derived. The section ends with a discussion of dispersion and damping of transversal waves propagating through a plasma, leading to the definition (4.115) of the dispersion measure.

4.4.1 Longitudinal and transversal dielectricities

In a thermal plasma with temperature T, the equilibrium phase-space distribution f_0 of the electrons can be assumed to be a Maxwellian,

$$f_0(\vec{p}) = \frac{\bar{n}}{(2\pi\sigma^2)^{3/2}} \mathrm{e}^{-p^2/(2\sigma^2)} , \quad \sigma = \sqrt{m_\mathrm{e} k_\mathrm{B} T} . \tag{4.88}$$

With this specific choice, the integrals in the longitudinal and the transversal dielectricities, (4.69) and (4.70) respectively, can be worked out.

Without loss of generality, we can first conveniently rotate the coordinate frame such that $\vec{k}$ points into the positive $\vec{x}$ direction. Then, $\vec{v}_\perp$ falls into the y-z plane. Since the derivative of f_0 with respect to any momentum component p_i is

$$\frac{\partial f_0(\vec{p})}{\partial p_i} = -\frac{p_i}{\sigma^2} f_0(\vec{p}) , \tag{4.89}$$

the remaining integrals in (4.69) and (4.70) read

$$\int \mathrm{d}^3 p\, \vec{k} \cdot \frac{\partial f_0}{\partial \vec{p}} \frac{1}{\vec{k} \cdot \vec{v} - \omega} = -\int \mathrm{d}^3 p\, \frac{k p_x}{\sigma^2} f_0(\vec{p}) \frac{1}{k p_x/m - \omega} \tag{4.90}$$

and

$$\int \mathrm{d}^3 p\, \vec{v}_\perp \cdot \frac{\partial f_0}{\partial \vec{p}_\perp} \frac{1}{\vec{k} \cdot \vec{v} - \omega} = -\frac{1}{m} \int \mathrm{d}^3 p\, \frac{p_y^2 + p_z^2}{\sigma^2} f_0(\vec{p}) \cdot \frac{1}{k p_x/m - \omega} . \tag{4.91}$$

The integrations over p_y and p_z in (4.91) and (4.91) can immediately be carried out, resulting in

$$\int \mathrm{d}^3 p\, \vec{k} \cdot \frac{\partial f_0}{\partial \vec{p}} \frac{1}{\vec{k} \cdot \vec{v} - \omega} = -\frac{\bar{n}}{(2\pi\sigma^2)^{1/2}} \frac{k}{\sigma^2} \int_{-\infty}^{\infty} \mathrm{d}p_x\, \frac{p_x\, \mathrm{e}^{-p_x^2/(2\sigma^2)}}{k p_x/m - \omega} \tag{4.92}$$

and

$$\int \mathrm{d}^3 p\, \vec{v}_\perp \cdot \frac{\partial f_0}{\partial \vec{p}_\perp} \frac{1}{\vec{k} \cdot \vec{v} - \omega} = -\frac{2\bar{n}}{(2\pi\sigma^2)^{1/2} m} \int_{-\infty}^{\infty} \mathrm{d}p_x\, \frac{\mathrm{e}^{-p_x^2/(2\sigma^2)}}{k p_x/m - \omega} . \tag{4.93}$$

To simplify the remaining integrals in (4.92) and (4.93), we introduce the dimension-less momentum component

$$t := \frac{p_x}{\sqrt{2}\sigma} \tag{4.94}$$

and the dimension-less frequency

$$y = \frac{m\omega}{\sqrt{2}k\sigma} = \frac{\omega}{\omega_\mathrm{th}} \quad \text{with} \quad \omega_\mathrm{th} := \frac{\sqrt{2}k\sigma}{m} = k\sqrt{\frac{2k_\mathrm{B}T}{m}} \tag{4.95}$$

to bring (4.92) and (4.93) into the forms

$$\int \mathrm{d}^3p\,\vec{k}\cdot\frac{\partial f_0}{\partial\vec{p}}\,\frac{1}{\vec{k}\cdot\vec{v}-\omega} = \frac{\bar{n}}{(2\pi\sigma^2)^{1/2}}\,\frac{\sqrt{2}m}{\sigma}\int_{-\infty}^{\infty}\mathrm{d}t\,\frac{t\,\mathrm{e}^{-t^2}}{y-t} \tag{4.96}$$

and

$$\int \mathrm{d}^3p\,\vec{v}_\perp\cdot\frac{\partial f_0}{\partial\vec{p}_\perp}\,\frac{1}{\vec{k}\cdot\vec{v}-\omega} = \frac{2\bar{n}}{(2\pi\sigma^2)^{1/2}k}\int_{-\infty}^{\infty}\mathrm{d}t\,\frac{\mathrm{e}^{-t^2}}{y-t}\;. \tag{4.97}$$

?

Since the integrands in (4.96) and (4.97) are singular at $t = y$, it remains to be shown that the integrals exist at all. How could you achieve this?

Before we can gain further insight into expressions like (4.96) and (4.97), we need to carefully evaluate the integrals appearing there because the integrands have a pole on the t axis at $t = y$. These integrals can be solved by continuing the integrand into the complex plane, $t \to z \in \mathbb{C}$, and then using the residue theorem from the theory of complex functions. Before doing so, we rewrite the integral in (4.96) as

$$\int_{-\infty}^{\infty}\mathrm{d}z\,\frac{z\,\mathrm{e}^{-z^2}}{y-z} = -\int_{-\infty}^{\infty}\mathrm{d}z\,\frac{(z-y+y)\,\mathrm{e}^{-z^2}}{z-y} = -\sqrt{\pi} + y\int_{-\infty}^{\infty}\frac{\mathrm{d}z\,\mathrm{e}^{-z^2}}{y-z}\;. \tag{4.98}$$

The remaining integral is given by the so-called *Faddeeva function* $w(z)$ which, for a positive imaginary part of z, has the integral representation

$$w(z) = \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-t^2}}{z-t}\mathrm{d}t\;. \tag{4.99}$$

If we choose the imaginary part of y to be arbitrarily small, we may thus express the integral from (4.96) by

$$\int_{-\infty}^{\infty}\mathrm{d}t\,\frac{t\,\mathrm{e}^{-t^2}}{y-t} = -\sqrt{\pi} - \mathrm{i}\pi y\,w(y) = -\sqrt{\pi}\,[1 + y\,Z(y)] \quad\text{with}\quad Z(y) := \mathrm{i}\sqrt{\pi}w(y)\;. \tag{4.100}$$

?

Based on the definitions of the Debye wavelength and the plasma frequency, confirm the expressions (4.101) and (4.102) for $\hat{\varepsilon}_\parallel$ and $\hat{\varepsilon}_\perp$.

The function $Z(y)$ defined here is called *plasma dispersion function*.

Returning with (4.100) into (4.96) and inserting the result into (4.69), we can now write the longitudinal dielectricity (4.69) as

$$\hat{\varepsilon}_\parallel = 1 + \frac{1 + y\,Z(y)}{\lambda_\mathrm{D}^2k^2} = 1 + \frac{\omega_\mathrm{p}^2}{\omega^2}\,2y^2\,[1 + y\,Z(y)]\;, \tag{4.101}$$

where the Debye wavelength λ_D as defined in (4.11) was identified in the first step and the plasma frequency ω_p from (4.16) as well as the dimension-less frequency y from (4.95) in the second. Similarly, identifying the Faddeeva function (4.99) in (4.97), using the definition (4.100) of the plasma dispersion function and inserting these expressions into (4.70), we find the simple expression

$$\hat{\varepsilon}_\perp = 1 + \frac{\omega_\mathrm{p}^2}{\omega^2}\,y\,Z(y) \tag{4.102}$$

for the transversal dielectricity. Here, we have again inserted the plasma frequency ω_p defined in (4.16) and used the definition (4.95) of the dimension-less frequency y. We have thus reduced the longitudinal and transverse dielectricities of a thermal plasma to the plasma dispersion function $Z(y)$, with the scaled frequency y defined by (4.95).

Series expansions of $Z(y)$ are useful for practical calculations (Figure 4.4). For small $|y| \ll 1$,

$$y\,Z(y) \approx -2y^2\left(1 - \frac{2y^2}{3} + \frac{4x^4}{15} - \dots\right) + \mathrm{i}\sqrt{\pi}\,y\left(1 - y^2\right) \tag{4.103}$$

while for large $|y| \gg 1$,

$$y\,Z(y) \approx -1 - \frac{1}{2y^2} - \frac{3}{4y^4} + \dots\ . \tag{4.104}$$

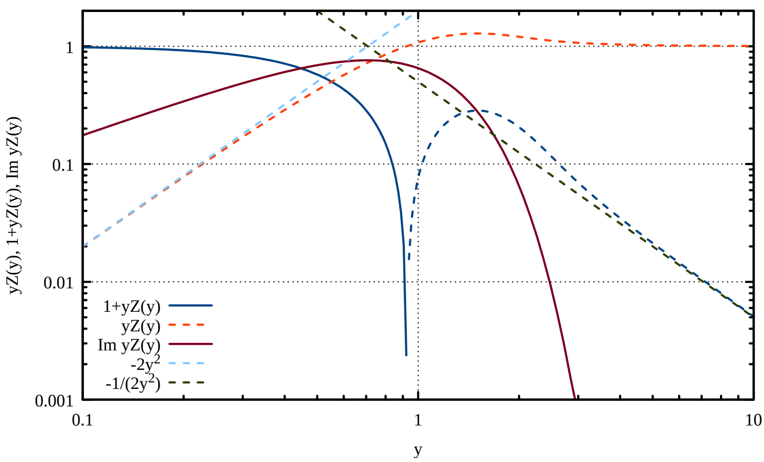

Figure 4.4 The absolute value of the plasma dispersion function $Z(y)$ as a function of y is plotted along the real axis, together with the approximations given in (4.103) and (4.104).

Before we begin discuss the results (4.101) and (4.102) in more detail, we recall the definition (4.95) of y and introduce the plasma frequency ω_p from (4.16) and the Debye wave number k_D from (4.6) into it,

$$y = \frac{\omega m}{\sqrt{2}k\sigma} = \frac{\omega}{\omega_\mathrm{p}}\frac{k_\mathrm{D}}{k}\frac{\omega_\mathrm{p}}{k_\mathrm{D}}\frac{m}{\sqrt{2}\sigma} = \frac{\tilde{\omega}}{\sqrt{2}\tilde{k}}\ , \tag{4.105}$$

inserting the definition (4.88) of σ in the final step. Quantities with a tilde now refer to plasma units,

$$\tilde{\omega} = \frac{\omega}{\omega_\mathrm{p}}\ , \quad \tilde{k} = \frac{k}{k_\mathrm{D}}\ . \tag{4.106}$$

In terms of these quantities, we can write the dielectricities (4.101) and (4.102) as

$$\hat{\varepsilon}_\parallel = 1 + \frac{1 + y\,Z(y)}{\tilde{k}^2}\ , \quad \hat{\varepsilon}_\perp = 1 + \frac{y\,Z(y)}{\tilde{\omega}^2}\ . \tag{4.107}$$

With the series expansions (4.103) and (4.104) of the plasma dispersion function, we arrive at the formal expressions for $\hat{\varepsilon}_\parallel$ and $\hat{\varepsilon}_\perp$ listed in Tab. 4.1.

?

Verify the entries in Tab. 4.1.

If multiple particle species need to be taken into account in addition to the electrons, the individual dielectricities are summed over all species according to

$$\hat{\varepsilon} - 1 = \sum_{\text{species}\,i} (\hat{\varepsilon}_i - 1)\ . \tag{4.108}$$

Table 4.1 Limiting cases for the longitudinal and transverse dielectricities, $\hat{\varepsilon}_{\|}$ and $\hat{\varepsilon}_{\perp}$, for small and large values of $|y|$.

dielectricity	$\|y\| \ll 1,\ \tilde{\omega} \ll \tilde{k}$	$\|y\| \gg 1,\ \tilde{k} \ll \tilde{\omega}$
$\hat{\varepsilon}_{\|}$	$1+\frac{1}{\tilde{k}^2}\left(1-\frac{\tilde{\omega}^2}{\tilde{k}^2}\right)+\mathrm{i}\sqrt{\frac{\pi}{2}}\frac{\tilde{\omega}}{\tilde{k}^3}$	$1-\frac{1}{\tilde{\omega}^2}\left(1+\frac{3\tilde{k}^2}{\tilde{\omega}^2}\right)$
$\hat{\varepsilon}_{\perp}$	$1-\frac{1}{\tilde{k}^2}\left(1-\frac{1}{3}\frac{\tilde{\omega}^2}{\tilde{k}^2}\right)+\mathrm{i}\sqrt{\frac{\pi}{2}}\frac{1}{\tilde{k}\tilde{\omega}}$	$1-\frac{1}{\tilde{\omega}^2}\left(1+\frac{\tilde{k}^2}{\tilde{\omega}^2}\right)$

4.4.2 Dispersion Measure and Dispersion Relations

The dispersion relation for transversal waves was given by (4.53). We combine it with the relation

$$\omega_{\mathrm{p}}^2 = \frac{k_{\mathrm{B}}T}{m}k_{\mathrm{D}}^2 = \left\langle v^2 \right\rangle k_{\mathrm{D}}^2\ . \tag{4.109}$$

between the plasma frequency and the Debye wavenumber and introduce the root-mean square velocity from (4.14). Defining a mean-squared beta factor by

$$\beta^2 := \frac{1}{c^2}\left\langle v^2 \right\rangle\ , \tag{4.110}$$

we can write the dispersion relation (4.53) in dimension-less form as

$$\beta^2\tilde{\omega}^2 = \frac{\tilde{k}^2}{\hat{\varepsilon}_{\perp}} \tag{4.111}$$

in plasma units. For high frequencies $\tilde{\omega} \gg \tilde{k}$, we can approximate the transverse dielectricity fom Tab. 4.1 by $\hat{\varepsilon}_{\perp} \approx 1 - \tilde{\omega}^{-2}$ and find the dispersion relation

$$\tilde{\omega}^2 = \frac{\tilde{k}^2}{\beta^2} + 1 \quad \text{or} \quad \tilde{k} = \beta\sqrt{\tilde{\omega}^2 - 1}\ . \tag{4.112}$$

Since the group velocity of such transversal waves is

$$c_{\mathrm{g}} = \frac{\partial\omega}{\partial k} = \frac{\omega_{\mathrm{p}}}{k_{\mathrm{D}}}\frac{\partial\tilde{\omega}}{\partial\tilde{k}} = \frac{c}{\beta}\frac{\tilde{k}}{\tilde{\omega}} = c\sqrt{1-\frac{1}{\tilde{\omega}^2}}\ , \tag{4.113}$$

the propagation time of such waves is

$$\Delta t_{\omega} = \int\frac{\mathrm{d}l}{c_{\mathrm{g}}} \approx \int\frac{\mathrm{d}l}{c}\left(1+\frac{\omega_{\mathrm{p}}^2}{2\omega^2}\right) = \frac{L}{c} + \frac{2\pi e^2}{mc\omega^2}\int\mathrm{d}l\,n\ , \tag{4.114}$$

where the integral over the electron density along the light path,

$$\boxed{\int\mathrm{d}l\,n \equiv \mathrm{DM}}\ , \tag{4.115}$$

is called the *dispersion measure*. In the last steps, we have used $\omega \gg \omega_{\mathrm{p}}$ to approximate the square root in the group velocity (4.113).

For lower frequencies, we need to take one more order in $\tilde{k}/\tilde{\omega}$ into account. Taking the expression for $\hat{\varepsilon}_{\perp}$ from the second row and the second column of

Tab. 4.1 and inserting it into the dimension-less dispersion relation (4.111), we find first

$$\beta^2\tilde{\omega}^2\left[1-\frac{1}{\tilde{\omega}^2}\left(1+\frac{\tilde{k}^2}{\tilde{\omega}^2}\right)\right]=\tilde{k}^2 \tag{4.116}$$

or, solving for the wavenumber $\tilde{k}$,

$$\tilde{k}=\beta\tilde{\omega}\sqrt{\frac{\tilde{\omega}^2-1}{\tilde{\omega}^2+\beta^2}}\,. \tag{4.117}$$

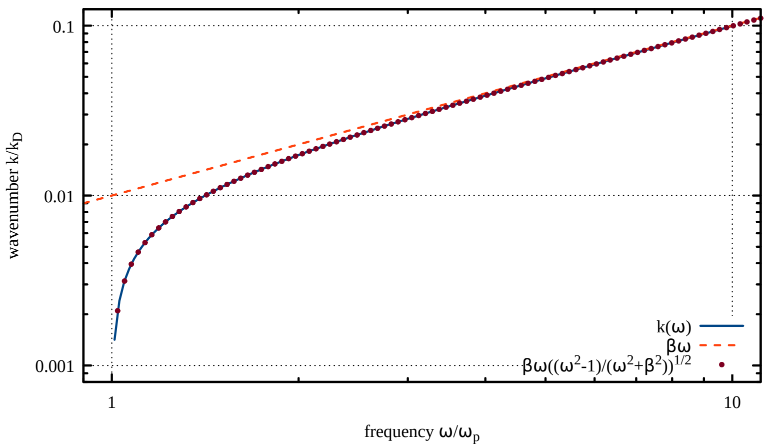

Figure 4.5 The dispersion relation for transverse waves is shown together with the approximations $\tilde{k}=\beta\tilde{\omega}$ and (4.117).

?

Solve the dispersion relation (4.116) for the frequency $\tilde{\omega}$.

Example: Ionosphere of the Earth

In the ionosphere of the Earth, $n \approx 10^6\,\mathrm{cm}^{-3}$. Assuming a temperature of $T = 273\,\mathrm{K}$, the Debye length and the root-mean square thermal electron velocity are

$$\lambda_\mathrm{D}\approx 0.11\,\mathrm{cm}\quad\text{and}\quad v_\mathrm{e}=\sqrt{\frac{k_\mathrm{B}T}{m}}\approx 6.43\cdot 10^6\,\mathrm{cm\,s^{-1}}\;; \tag{4.118}$$

see also (4.14). Thus, the Debye wavenumber is $k_\mathrm{D} = 9.1\,\mathrm{cm}^{-1}$, the β factor of the electrons is $\beta = 2.1\cdot 10^{-4}$, and the plasma frequency is

$$\omega_\mathrm{p}=\frac{v_\mathrm{e}}{\lambda_\mathrm{D}}\approx 58.5\,\mathrm{MHz}\,. \tag{4.119}$$

Transversal Gigahertz waves, for example, have $\tilde{\omega} \gg 1$, thus a transversal dielectricity of $\hat{\varepsilon}_\perp \approx 1$ and wavelengths of

$$\lambda=\frac{2\pi}{k}\approx 2\pi\frac{c}{\omega} \tag{4.120}$$

like in vacuum. Approaching the plasma frequency from above, the wavenumber falls below its vacuum value, and thus the waves become longer than in vacuum. ◀

The dispersion relation for longitudinal waves requires $\hat{\varepsilon}_{\parallel} = 0$, as was shown in (4.54) above. Assuming high frequencies,

the entry in the first row and the second column in Tab. 4.1 gives

$$\tilde{\omega}^2 = 1 + \frac{3\tilde{k}^2}{\tilde{\omega}^2} \tag{4.121}$$

or, solving for $\tilde{\omega}$,

$$\tilde{\omega}_{\pm}^2 = \frac{1}{2}\left(1 \pm \sqrt{1 + 12\tilde{k}^2}\right) . \tag{4.122}$$

Only the positive branch is meaningful here. In the high-frequency limit applied, we require $\tilde{\omega} \gg \tilde{k}$, thus $\tilde{\omega} \approx 1$ and $\tilde{k} \ll 1$. This allows us to approximate

$$\tilde{\omega} \approx 1 + \frac{3}{2}\tilde{k}^2 . \tag{4.123}$$

Such longitudinal waves thus have frequencies slightly higher than the plasma frequency and very large wavelengths.

Problems

1. Derive the phase and the group velocities of longitudinal and transverse electromagnetic waves in a thermal plasma in the high-frequency limit.

2. Derive the series expansions (4.103) and (4.104) of the plasma dispersion function in the limits $y \ll 1$ and $y \gg 1$.

3. Radio waves propagating through a thermal plasma in a deep gravitational well experience two kinds of time delay: the delay (4.114) due to the electromagnetic dispersion and the so-called Shapiro delay

$$\Delta t_{\text{Shapiro}} = -\frac{2}{c^3}\int \mathrm{d}l\,\Phi , \tag{4.124}$$

due to generally-relativistic time dilation, where Φ is the Newtonian gravitational potential. Estimate the relative magnitude of both time delays.

4.5 The Magnetohydrodynamic Equations

In this section, we introduce the assumptions of magnetohydrodynamics and derive the induction equation (4.141) for the evolution of the magnetic field in a plasma. Magnetic forces on the plasma appear in the extension (4.145) of Euler's equation, and we show in (4.161) how the energy current density has to be modified in presence of a magnetic field. The comparison of magnetic advection and diffusion leads to the definition of the magnetic Reynolds number in (4.169).

4.5.1 Assumptions

Magnetohydrodynamics is the theory of how magnetised plasmas move. It is built upon several assumptions which go significantly beyond hydrodynamics. They begin with the fact that plasmas consist of ions and electrons which should in the simplest case be described as two fluids coupled to each other rather than a single fluid, as in hydrodynamics. At this point, recall that the constitutive assumption of hydrodynamics was that the mean-free path of the fluid particles is very small compared to all other relevant length scales occurring in the system. In ideal hydrodynamics, the mean-free path is infinitely short. Giving up this idealisation, but still assuming that the mean-free path is very short, gives rise to effects based on particle transport, such as viscosity and diffusion.

The common, greatly simplifying assumption in magnetohydrodynamics is that the ions and the electrons are so tightly coupled to each other by their electrodynamic interaction that they can indeed be treated as a single fluid. The central assumption of hydrodynamics is then applied in addition, namely that the mean-free path of the plasma particles, ions and electrons alike, is very small.

If, however, there is no net motion of the electrons with respect to the ions, then there is no separation of charges, no net electric current, and thus neither an electric nor a magnetic field. For magnetohydrodynamics, therefore, we need to assume that there is in fact a small drift velocity $\vec{v}_{\rm drift}$ between the electrons and ions,

$$\vec{v}_{\rm drift} = \vec{v}_{\rm e} - \vec{v}_{\rm i} \; , \tag{4.125}$$

causing an electric current $\vec{j}$ of free charges, which can sustain a magnetic field. As in non-ideal, viscous hydrodynamics, the strict idealisation of two fluids of opposite charge infinitely tightly coupled to each other is slightly loosened here.

A final, common assumption is that the plasma flows non-relativistically, allowing us to neglect terms of higher than linear order in v/c, where v is the flow velocity.

More quantitatively, we thus arrive at the following assumptions: We first transform into the rest frame of the plasma, defined as the frame locally co-moving with the mean velocity of the two or more plasma components. Quantities in this rest frame are primed.

The plasma is macroscopically neutral, but as it consists of charged particles, even small drift velocities can create substantial currents. This is expressed by supposing that the time component of the current-density four-vector j'^{μ} observed in this rest frame be negligibly small compared to the spatial components of the current density,

$$c\rho' \ll \left|\vec{j}'\right| \; . \tag{4.126}$$

As usual in electrodynamics, the spatial current density $\vec{j}'$ itself is assumed to be related to the electric field $\vec{E}'$ through Ohm's law by the conductivity σ,

$$\vec{j}' = \sigma\vec{E}' \; . \tag{4.127}$$

The plasma is assumed to be an ideal or near-ideal conductor such that even weak electric fields can be responsible for significant currents. By its definition, the conductivity must have the dimension

$$[\sigma] = \text{time}^{-1} \ . \tag{4.128}$$

Accordingly, very high conductivity means that the time scale needed by the plasma charges to respond to changes in the electromagnetic fields is very small. This allows us to neglect the displacement current compared to the charge current,

$$\frac{\partial \vec{E}'}{\partial t} \ll \vec{j}' \ , \tag{4.129}$$

because any change in the electric field will immediately (or at least on a very short time scale of order σ^{-1}) lead to a substantial charge current. Maxwell's equations for the magnetic field thus simplify to read

$$\vec{\nabla} \cdot \vec{B}' = 0 \ , \quad \vec{\nabla} \times \vec{B}' = \frac{4\pi}{c}\vec{j}' \tag{4.130}$$

in the rest frame of the plasma. This rest frame and the observer's laboratory frame are related by a Lorentz transform as given in (1.26). However, by the assumption of non-relativistic plasma flow relative to the laboratory frame, we can expand the Lorentz factor γ to lowest order in $\beta = v/c$, i.e. we can adopt $\gamma \approx 1$. Then, the Lorentz transform of the four-current gives

$$c\rho' = c\rho - \vec{\beta} \cdot \vec{j} \ , \quad \vec{j}' = \vec{j} - \vec{\beta} c\rho \ . \tag{4.131}$$

Since we have assumed $c\rho' \ll |\vec{j}'|$ in the plasma's rest frame, we also have $c\rho \ll |\vec{j}|$ in the laboratory frame due to the non-relativistic plasma flow, $\beta \ll 1$. Since we must further obey the Maxwell equations

$$\vec{\nabla} \cdot \vec{E}' = 4\pi\rho' \ , \quad \vec{\nabla} \times \vec{B}' = \frac{4\pi}{c}\vec{j}' \ , \tag{4.132}$$

this also implies $|\vec{E}| \ll |\vec{B}|$. Accordingly, the assumptions of (non-relativistic) magnetohydrodynamics imply the conditions

$$c\rho \ll \left|\vec{j}\right| \ , \quad \left|\vec{E}\right| \ll \left|\vec{B}\right| \ , \quad \left|\frac{\partial \vec{E}}{\partial t}\right| \ll \left|\vec{j}\right| \ , \quad \beta = \frac{\left|\vec{v}\right|}{c} \ll 1 \ . \tag{4.133}$$

?

Confirm expressions (4.131) by carrying out a Lorentz transform in the appropriate limit.

On the basis of these relations, we can now derive the equations of magnetohydrodynamics.

4.5.2 The induction equation

We begin with Ohm's law

$$\vec{j} \approx \vec{j}' = \sigma \vec{E}' \tag{4.134}$$

and relate the electric field $\vec{E}'$ in the plasma's rest frame to the electric field $\vec{E}$ in the laboratory frame by the Lorentz transform (1.87) in the limit $\gamma \approx 1$. We can thus insert

$$\vec{E}' = \vec{E} + \vec{\beta} \times \vec{B} \tag{4.135}$$

into (4.134) and to solve for $\vec{E}$ to obtain

$$\vec{E} = \frac{\vec{j}}{\sigma} - \vec{\beta} \times \vec{B} \,. \tag{4.136}$$

Next, we put this result into the induction equation

$$\frac{\partial \vec{B}}{\partial t} = -c\vec{\nabla} \times \vec{E} \tag{4.137}$$

and find

$$\frac{\partial \vec{B}}{\partial t} = -c\vec{\nabla} \times \left(\frac{\vec{j}}{\sigma} - \vec{\beta} \times \vec{B} \right) \tag{4.138}$$

for the evolution of the magnetic field. At the same time, we need to satisfy Ampère's law with vanishing displacement current,

$$\vec{j} = \frac{c}{4\pi} \vec{\nabla} \times \vec{B} \,, \tag{4.139}$$

which enables us to eliminate the current density from the induction equation. Using the identity

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{B} \right) = \vec{\nabla} \left(\vec{\nabla} \cdot \vec{B} \right) - \vec{\nabla}^2 \vec{B} \tag{4.140}$$

for the double curl and Maxwell's equation $\vec{\nabla} \cdot \vec{B} = 0$ for the divergence of the magnetic field, we find

$$\frac{\partial \vec{B}}{\partial t} = \frac{c^2}{4\pi\sigma} \vec{\nabla}^2 \vec{B} + \vec{\nabla} \times \left(\vec{v} \times \vec{B} \right) \tag{4.141}$$

if we further assume that the conductivity is spatially constant, $\vec{\nabla}\sigma = 0$. This *induction equation* determines the evolution of the magnetic field embedded into a plasma flow with the velocity $\vec{v}$.

?

What do the two terms on the right-hand side of the induction equation (4.141) mean, i.e. what physical effects do they encode?

4.5.3 Euler's equation

The induction equation tells us how the magnetic field changes in response to the plasma flow. In addition, we need equations for the back-reaction of the magnetic field on the plasma flow. We have to expect that a magnetised plasma flows differently than a neutral fluid because the magnetic field acts on the charged particles through the Lorentz force.

Notice, however, that the continuity equation for the mass density ρ will remain unchanged,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \,, \tag{4.142}$$

because mass conservation must not be affected by the presence of the magnetic field. Euler's equation, which describes the conservation of momentum or, more precisely, the transport of the specific momentum density, must be modified by the presence of the Lorentz force. In absence of electric fields, the force exerted on a charge e by the magnetic field $\vec{B}$ is

$$\frac{e}{c} \vec{v} \times \vec{B} \,. \tag{4.143}$$

Multiplying this expression with the number density n of charges will turn it into the Lorentz force density, i.e. into the combined Lorentz force on all charges in a unit volume. Noticing that the product $ne\vec{v}$ is the current density $\vec{j}$, and eliminating the current density once more by Ampère's law (4.139), we find that the magnetic force density on the plasma must be

$$\frac{1}{4\pi}\left(\vec{\nabla}\times\vec{B}\right)\times\vec{B}\,, \tag{4.144}$$

and this term must be added to Euler's equation, which now reads

$$\rho\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\vec{\nabla}P + \frac{1}{4\pi}\left(\vec{\nabla}\times\vec{B}\right)\times\vec{B} \tag{4.145}$$

in absence of gravitational forces. By means of the identity

$$\left(\vec{\nabla}\times\vec{B}\right)\times\vec{B} = \left(\vec{B}\cdot\vec{\nabla}\right)\vec{B} - \frac{1}{2}\vec{\nabla}\left(\vec{B}^2\right)\,, \tag{4.146}$$

the Lorentz force density acting on the plasma can be cast into the very intuitive form

$$\frac{1}{4\pi}\left(\vec{\nabla}\times\vec{B}\right)\times\vec{B} = \frac{1}{4\pi}\left(\vec{B}\cdot\vec{\nabla}\right)\vec{B} - \frac{1}{8\pi}\vec{\nabla}\left(\vec{B}^2\right)\,. \tag{4.147}$$

The first term specifies how $\vec{B}$ changes along $\vec{B}$, i.e. it quantifies the tension of the magnetic field lines, which obviously tend to be as straight as possible. The second term is the gradient of the magnetic energy density and augments the pressure gradient in Euler's equation. We thus find that the magnetic field acts in two ways on the plasma flow: It resists motions that bend and compress the field (Figure 4.6).

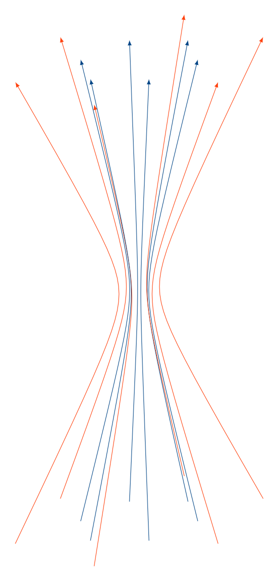

Figure 4.6 A magnetic field exert two kinds of force on a plasma: The field lines tend to straighten due to the $(\vec{B}\cdot\vec{\nabla})\vec{B}$ term, and they tend to reduce the magnetic pressure due to the $\vec{\nabla}(\vec{B}^2)$ term in the magnetic Euler equation.

We have seen in normal, viscous hydrodynamics that Euler's equation can be written in the manifestly conservative form

$$\partial_t (\rho\vec{v}) + \vec{\nabla} \cdot \bar{T} = 0 \,, \tag{4.148}$$

where the stress-energy tensor

$$\bar{T} = \rho\vec{v} \otimes \vec{v} + P\mathbb{1}_3 + \bar{T}_\mathrm{d} \tag{4.149}$$

occurs. It contains the diffusive contribution T_d, given in (3.143), which reads

$$\bar{T}_\mathrm{d} = -\eta \left[\left(\vec{\nabla} \otimes \vec{v}\right) + \left(\vec{\nabla} \otimes \vec{v}\right)^\top - \frac{2}{3}\vec{\nabla} \cdot \vec{v}\,\mathbb{1}_3 \right] - \zeta\vec{\nabla} \cdot \vec{v}\,\mathbb{1}_3 \,. \tag{4.150}$$

In presence of a magnetic field, the stress-energy tensor must be augmented by a magnetic contribution,

$$\bar{T} \to \bar{T} + \bar{T}_\mathrm{m} \,, \tag{4.151}$$

whose components are given in (1.111),

$$\bar{T}_\mathrm{m} = -\frac{1}{4\pi} \left(\vec{B} \otimes \vec{B} - \frac{\vec{B}^2}{2}\,\mathbb{1}_3 \right) . \tag{4.152}$$

The stress-energy tensor of the ideal fluid (3.51), a the diffusive part (3.143) for the viscous fluid and the contribution by the magnetic field (4.152) are thus simply added.

Together with an equation of state, $P = P(\rho)$, the induction equation (4.141), the continuity equation (4.142) and Euler's equation (4.145) determine both the plasma flow and the evolution of the magnetic field embedded into it. These are two scalar and two vector equations, thus eight equations, for the eight unknowns ρ, P, $\vec{v}$ and $\vec{B}$. If the magnetic field is known, the current follows from Ampère's law (4.139), and the electric field is finally given by (4.136).

4.5.4 Energy and entropy

The evolution equation (3.163) for the specific entropy $\tilde{s}$ per unit mass, which read

$$\rho T \frac{\mathrm{d}\tilde{s}}{\mathrm{d}t} = \vec{\nabla} \cdot \left(\kappa\vec{\nabla}T\right) + \mathrm{Tr}\left(\bar{T}_\mathrm{d} Dv\right) \,, \tag{4.153}$$

in ordinary, viscous hydrodynamics of neutral fluids, now needs to be augmented by the entropy production through the release of Ohmic heat.

?

Entropy production by which physical process does the new term on the right-hand side of (4.155) describe?

Per unit time, the induction current $\vec{j}'$ in the rest frame of the fluid dissipates the energy

$$\vec{j}' \cdot \vec{E}' = \frac{\vec{j}'^2}{\sigma} \approx \frac{\vec{j}^2}{\sigma} = \frac{c^2}{16\pi^2\sigma} \left(\vec{\nabla} \times \vec{B}\right)^2 \,, \tag{4.154}$$

where we have used Ohm's law in the first, $\vec{j}' \approx \vec{j}$ in the second and Ampère's law (4.139) in the last steps. The resulting expression must be added to the right-hand side of the entropy equation, giving

$$\rho T \frac{\mathrm{d}\tilde{s}}{\mathrm{d}t} = \vec{\nabla} \cdot \left(\kappa\vec{\nabla}T\right) + \mathrm{Tr}\left(\bar{T}_\mathrm{d} Dv\right) + \frac{c^2}{16\pi^2\sigma} \left(\vec{\nabla} \times \vec{B}\right)^2 \,. \tag{4.155}$$

If we need to express energy conservation by the specific energy density ε instead of the specific entropy density s, we start from the energy conservation equation of viscous hydrodynamics and augment it in a completely analogous way. First, the energy density of the magnetic field, $\vec{B}^2/(8\pi)$, must be added to the kinetic and thermal energy density of the fluid,

$$\frac{\rho}{2}\vec{v}^2 + \varepsilon \to \frac{\rho}{2}\vec{v}^2 + \varepsilon + \frac{\vec{B}^2}{8\pi} \,. \tag{4.156}$$

Next, the Poynting vector of the electromagnetic field,

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} \,, \tag{4.157}$$

must be added to the energy current density $\vec{q}$. In ordinary, viscous hydrodynamics, its components were given by (3.51) and (3.151)

$$\vec{q} = \rho\left(\frac{\vec{v}^2}{2} + \tilde{h}\right)\vec{v} - \kappa\vec{\nabla}T - \bar{T}_{\rm d}\vec{v} \,. \tag{4.158}$$

Using (4.136) and Ampère's law (4.139) once more, the Poyting vector can be expressed by

$$\vec{S} = \frac{c^2}{16\pi^2\sigma}\left(\vec{\nabla}\times\vec{B}\right)\times\vec{B} - \frac{1}{4\pi}\left(\vec{v}\times\vec{B}\right)\times\vec{B} \,, \tag{4.159}$$

which we can rearrange by expanding the vector products into

$$\vec{S} = \frac{c^2}{16\pi^2\sigma}\left[\left(\vec{B}\cdot\vec{\nabla}\right)\vec{B} - \frac{1}{2}\vec{\nabla}\vec{B}^2\right] - \frac{1}{4\pi}\left[\left(\vec{B}\cdot\vec{v}\right)\vec{B} - \vec{B}^2\,\vec{v}\right] \tag{4.160}$$

Thus, the energy current density in a viscous, magnetised plasma has the components

$$\begin{aligned}\vec{q} &= \rho\left(\frac{\vec{v}^2}{2} + \tilde{h}\right)\vec{v} - \kappa\vec{\nabla}T - \bar{T}_{\rm d}\vec{v} \\ &- \frac{c^2}{16\pi^2\sigma}\left[\left(\vec{B}\cdot\vec{\nabla}\right)\vec{B} - \frac{1}{2}\vec{\nabla}\vec{B}^2\right] - \frac{1}{4\pi}\left[\left(\vec{B}\cdot\vec{v}\right)\vec{B} - \vec{B}^2\,\vec{v}\right] \,.\end{aligned} \tag{4.161}$$

Each of these terms has an intuitive physical meaning: The first term in parentheses is the transport of kinetic energy and enthalpy with the fluid flow, where the enthalpy appears instead of the kinetic energy to take any pressure-volume work into account that the fluid may have to exert. The following two terms describe energy loss by heat conduction and by viscous friction. The next term in brackets is multiplied with the inverse conductivity and thus disappears if the plasma is ideally conducting, $\sigma \to \infty$. The first term in brackets is the magnetic tension, the second is the gradient of the internal energy of the magnetic field. In the final bracket, the first term quantifies how the magnetic field changes along the flow lines of the fluid, and the final term is the transport of the magnetic energy with the fluid.

4.5.5 Incompressible flows

For incompressible flows with $\vec{\nabla} \cdot \vec{v} = 0$, the magnetohydrodynamic equations simplify somewhat. First, expanding the curl of the vector product in the induction equation (4.141) and using $\vec{\nabla} \cdot \vec{v} = 0$ in addition to $\vec{\nabla} \cdot \vec{B} = 0$, we find

$$\frac{\mathrm{d}\vec{B}}{\mathrm{d}t} = \frac{\partial \vec{B}}{\partial t} + \left(\vec{v} \cdot \vec{\nabla}\right) \vec{B} = \left(\vec{B} \cdot \vec{\nabla}\right) \vec{v} + \frac{c^2}{4\pi\sigma} \vec{\nabla}^2 \vec{B} \,, \tag{4.162}$$

and Euler's equation becomes

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \frac{\partial \vec{v}}{\partial t} + \left(\vec{v} \cdot \vec{\nabla}\right) \vec{v} = -\frac{1}{\rho} \vec{\nabla} \left(P + \frac{\vec{B}^2}{8\pi} \right) + \frac{1}{4\pi\rho} \left(\vec{B} \cdot \vec{\nabla}\right) \vec{B} + \nu \vec{\nabla}^2 \vec{v} \,, \tag{4.163}$$

where $\nu = \eta/\rho$ is the specific viscosity per unit mass. Moreover, the diffusive stress-energy tensor T_d in the energy-conservation equation simplifies to read

$$\bar{T}_\mathrm{d} = -\eta \left[\left(\vec{\nabla} \otimes \vec{v}\right) + \left(\vec{\nabla} \otimes \vec{v}\right)^\top \right] = -2\eta D v \,, \tag{4.164}$$

where the symmetrised velocity-gradient tensor Dv from (3.154) was inserted. This enables us to bring the viscous dissipation term in the energy-conservation equation into the simple form

$$\mathrm{Tr}\left[\bar{T}_\mathrm{d}^\top \left(\vec{\nabla} \otimes \vec{v}\right) \right] = \frac{1}{2} \mathrm{Tr}\left(\bar{T}_\mathrm{d}^\top D v \right) = -\eta \; \mathrm{Tr}\left(D v^\top D v \right) \,. \tag{4.165}$$

4.5.6 Magnetic advection and diffusion

Two terms determine the temporal change of the magnetic field in the induction equation (4.141),

$$\vec{\nabla} \times \left(\vec{v} \times \vec{B}\right) \quad \text{and} \quad \frac{c^2}{4\pi\sigma} \vec{\nabla}^2 \vec{B} \,. \tag{4.166}$$

The first term, $\vec{\nabla} \times (\vec{v} \times \vec{B})$, determines the transport of the magnetic field with the fluid flow. It is called *advection term*. Its order-of-magnitude is

$$\frac{vB}{L} \,, \tag{4.167}$$

if L is a typical length scale characterising the plasma flow. The second term, proportional to $\vec{\nabla}^2 \vec{B}$, determines the diffusion of the magnetic field due to the finite conductivity of the plasma. If the conductivity is ideally large, $\sigma \to \infty$, the diffusion coefficient vanishes, showing that magnetic fields cannot move with respect to an ideally conducting plasma.

?

Expand the curl of the vector product $\vec{v} \times \vec{B}$ and verify the meaning of this term noted in the text.

The diffusion term has the order of magnitude

$$\frac{c^2}{4\pi\sigma} \frac{B}{L^2} \,. \tag{4.168}$$

The order-of-magnitude ratio between the advection and diffusion terms,

$$\frac{\text{advection}}{\text{diffusion}} = \frac{4\pi\sigma}{c^2} \frac{L^2}{B} \frac{vB}{L} = \boxed{\frac{4\pi\sigma v L}{c^2} \equiv \mathcal{R}_\mathrm{M}} \,, \tag{4.169}$$

is called the *magnetic Reynolds number*. Obviously, the magnetic-field diffusion can be neglected if $\mathcal{R}_\mathrm{M} \gg 1$, and the induction equation simplifies to

$$\frac{\partial \vec{B}}{\partial t} - \vec{\nabla} \times \left(\vec{v} \times \vec{B}\right) = 0 \; . \tag{4.170}$$

In absence of diffusion, the magnetic field is said to be "frozen" into the plasma. The physical reason for this is that, if the conductivity is very high, $\sigma \to \infty$, each motion of the magnetic field with respect to the plasma immediately induces strong current densities which counter-act their origin, i.e. the motion of the field. This is a typical case in astrophysical plasmas.

In the opposite limit, $\mathcal{R}_\mathrm{M} \ll 1$, which occurs if the conductivity is small, the induction equation turns into the pure diffusion equation

$$\frac{\partial \vec{B}}{\partial t} = \frac{c^2}{4\pi\sigma} \vec{\nabla}^2 \vec{B} \; . \tag{4.171}$$

Transforming this equation into Fourier space immediately shows that Fourier modes of wave number k solve this equation if their frequency is

$$\omega_\mathrm{diff} = \mathrm{i}\frac{c^2 k^2}{4\pi\sigma} \; . \tag{4.172}$$

This means that magnetic field modes with wavelength $\lambda = 2\pi k^{-1}$ must decay exponentially on the diffusion time scale

$$\tau_\mathrm{diff} = \frac{2\pi}{\mathrm{Im}\,\omega_\mathrm{diff}} \approx 2\sigma \frac{\lambda^2}{c^2} \; , \tag{4.173}$$

which is directly proportional to the conductivity σ: The lower the conductivity is, the faster the magnetic field decays by diffusion. Plasmas in the laboratory are typically characterised by $\mathcal{R}_\mathrm{M} \ll 1$, while astrophysical plasmas typically have $\mathcal{R}_\mathrm{M} \gg 1$.

Problems

1. Specialise the expression (4.144) for the magnetic force density to the case of a magnetic field confined to the x-y plane or parallel to the z axis.
2. Consider a general diffusion equation for a function $f(t, \vec{x})$,

   $$\frac{\partial f}{\partial t} = C \vec{\nabla}^2 f \; , \tag{4.174}$$

   and show that it is solved by a convolution of an initial function $f_0(\vec{x})$ with a Gaussian. How does the width of this Gaussian evolve in time?

4.6 Generation of Magnetic Fields

In this section, we briefly touch the vast subject of how magnetic fields can be generated. We show how magnetic fields can be generated if electrons and ions are not ideally tightly coupled to each other and derive the modified induction equation (4.185) that now contains a source term given by any misalignment between the gradients of the electron pressure and the particle number density.

The induction equation (4.141) contains no source term: Both terms on its right-hand side, which together determine the time evolution of $\vec{B}$, are linear in $\vec{B}$. The equation can therefore only describe how existing magnetic fields change, but if $\vec{B} = 0$ initially, this remains so. This is a consequence of the assumption that ions and electrons are ideally (tightly) coupled to each other. Should this not be the case, the flows of the electrons and of the ions need to be considered separately, most notably with different velocities $\vec{v}_e$ and $\vec{v}_i$. Then, two separate Euler equations must hold for the electrons and the ions,

$$\begin{aligned} n_e m_e \frac{d\vec{v}_e}{dt} &= -\vec{\nabla} P_e - n_e e \left(\vec{E} + \frac{\vec{v}_e}{c} \times \vec{B} \right) - n_e m_e \vec{\nabla}\Phi \;, \\ n_i m_i \frac{d\vec{v}_i}{dt} &= -\vec{\nabla} P_i + n_i e \left(\vec{E} + \frac{\vec{v}_i}{c} \times \vec{B} \right) - n_i m_i \vec{\nabla}\Phi \;, \end{aligned} \tag{4.175}$$

which are coupled by common electromagnetic fields $\vec{E}$ and $\vec{B}$ and by the gravitational potential Φ. We divide these equations by $n_e m_e$ and $n_i m_i$, respectively, and subtract the second from the first to find the evolution of the relative velocity

$$\frac{d(\vec{v}_e - \vec{v}_i)}{dt} = -\frac{\vec{\nabla} P_e}{n_e m_e} + \frac{\vec{\nabla} P_i}{n_i m_i} - \frac{e}{m_e}\left(\vec{E} + \frac{\vec{v}_e}{c} \times \vec{B} \right) - \frac{e}{m_i}\left(\vec{E} + \frac{\vec{v}_i}{c} \times \vec{B} \right) \;. \tag{4.176}$$

Since the ion mass m_i is much larger than the electron mass m_e, but $n_e = n_i \equiv n$, equation (4.176) can be approximated by

$$\frac{d(\vec{v}_e - \vec{v}_i)}{dt} = -\frac{\vec{\nabla} P_e}{n m_e} - \frac{e}{m_e}\left(\vec{E} + \frac{\vec{v}_e}{c} \times \vec{B} \right) \;. \tag{4.177}$$

The terms containing the ion mass may be neglected because, if a relative velocity difference is to be maintained, it must be due to the lower inertia and thus the higher mobility of the electrons.

The last equation must be augmented by a phenomenological collision term through which different electron and ion velocities can be justified or produced in the first place. Introducing a collision time τ, we simply write

$$\frac{d(\vec{v}_e - \vec{v}_i)}{dt} = -\frac{\vec{\nabla} P_e}{n m_e} - \frac{e}{m_e}\left(\vec{E} + \frac{\vec{v}_e}{c} \times \vec{B} \right) - \frac{\vec{v}_e - \vec{v}_i}{\tau} \;. \tag{4.178}$$

The net current density of the electrons and the ions together is

$$\vec{j} = e n_i \vec{v}_i - e n_e \vec{v}_e = e n (\vec{v}_i - \vec{v}_e) \;, \tag{4.179}$$

where we have implicitly assumed singly-charged ions. This is not a severe restriction at all because whatever the ion charge Z is, we only need to make sure that the plasma is electrically neutral by satisfying (4.1).

For a stationary situation, the relative drift velocity between electrons and ions must be constant,

$$\frac{d(\vec{v}_i - \vec{v}_e)}{dt} = 0 \;, \tag{4.180}$$

which implies by (4.179) a constant total electric current density $\vec{j}$. In this situation, we can solve the drift equation (4.178) for the electric field $\vec{E}$ and eliminate the drift velocity by the current density $\vec{j}$ to find

$$\vec{E} = -\frac{\vec{\nabla} P_\mathrm{e}}{e n_\mathrm{e}} - \frac{\vec{v}_\mathrm{e}}{c} \times \vec{B} + \frac{m_\mathrm{e}\vec{j}}{n e^2 \tau} \ . \tag{4.181}$$

What does this electric field mean? If the last term on the right-hand side was missing, (4.181) would say that an equilibrium situation required an electric field which, when combined with the existing magnetic field, creates a Lorentz force balancing the pressure-gradient force. The phenomenological collision term on the right-hand side adds friction between the electrons and the ions. Seen from the perspective of the electrons, (4.181) now means that the particle collisions hinder the motion of the electrons and thereby enhance the electric field required for equilibrium.

By Faraday's law (1.97), this electric field determines the time evolution of the magnetic field,

$$\frac{\partial \vec{B}}{\partial t} = -c\vec{\nabla} \times \vec{E} = \frac{c}{e}\vec{\nabla} \times \frac{\vec{\nabla} P_\mathrm{e}}{n} + \vec{\nabla} \times \left(\vec{v}_\mathrm{e} \times \vec{B}\right) - \frac{m_\mathrm{e} c}{e^2 \tau}\vec{\nabla} \times \frac{\vec{j}}{n} \ . \tag{4.182}$$

Now, since the curl of the pressure gradient vanishes identically, $\vec{\nabla} \times \vec{\nabla} P_\mathrm{e}$, we can re-write the first term on the right-hand side as

$$\vec{\nabla} \times \frac{\vec{\nabla} P_\mathrm{e}}{n} = -\vec{\nabla} P_\mathrm{e} \times \frac{\vec{\nabla} n}{e n^2} \ . \tag{4.183}$$

?

In what kind of situations does the vector product $\vec{\nabla} P_\mathrm{e} \times \vec{\nabla} n$ between the pressure and number-density gradients not vanish? Construct examples.

The curl of the current density can be rewritten using Ampère's law, making use of the fundamental assumption of magnetohydrodynamics that displacement currents can be neglected; recall (4.182). We can then further conclude that

$$\vec{\nabla} \times \frac{\vec{j}}{n} = \frac{c}{4\pi n}\vec{\nabla} \times \left(\vec{\nabla} \times \vec{B}\right) + \vec{j} \times \frac{\vec{\nabla} n}{n^2} \ . \tag{4.184}$$

If the electric current is flowing along the gradient in the number density of the electrons, which is generally a reasonable assumption, the latter term vanishes identically, and we obtain the modified induction equation

$$\frac{\partial \vec{B}}{\partial t} = \frac{c^2}{4\pi\sigma}\vec{\nabla}^2 \vec{B} + \vec{\nabla} \times \left(\vec{v}_\mathrm{e} \times \vec{B}\right) - \frac{c}{e n^2}\left(\vec{\nabla} P_\mathrm{e} \times \vec{\nabla} n\right) \ , \tag{4.185}$$

where we have made use of the definition

$$\sigma \equiv \frac{n e^2 \tau}{m_\mathrm{e}} \tag{4.186}$$

of the conductivity. Compared to (4.141), this new induction equation is augmented by an inhomogeneity in terms of the magnetic field, given by the term on the right-hand side containing the gradient of the number density,

$$-\frac{c}{e n^2}\left(\vec{\nabla} P_\mathrm{e} \times \vec{\nabla} n\right) \ , \tag{4.187}$$

which now appears as a source of the magnetic field. It shows that magnetic fields can be created if there is a gradient in the number density of the electrons which is misaligned with the pressure gradient. Mechanisms like this for *creating* magnetic fields are called *battery mechanisms*.

Problems

1. The condition for an inhomogeneity of the form (4.187) to occur in the induction equation can be given in the simple form

$$\vec{\nabla} \times \left(\frac{\vec{\nabla} P}{\rho} \right) \neq 0 \,, \tag{4.188}$$

if the electron pressure and the electron number density are supposed to be proportional to the total gas pressure P and the gas density ρ.

   (a) Set up Euler's equation for an ideal fluid in hydrostatic equilibrium, ignoring magnetic forces, but adding the centrifugal force appearing if the fluid is rotating with an angular velocity $\vec{\Omega}$ about the z axis.
   (b) Can magnetic fields build up in a rotating object in hydrostatic equilibrium? If so, under which conditions on the rotation? How is the magnetic field oriented that is generated this way?
   (c) Suppose the rotation is uniform, but the fluid is chemically inhomogeneous such that the relation between the electron density and the matter density changes. Can magnetic fields be built up now?

4.7 Ambipolar Diffusion

This section discusses what happens if plasma and neutral gas are mixed: The neutral gas moves freely relative to the magnetic field, but remains coupled to the plasma by particle collisions. The first main result is the expression (4.208) for the density of the friction force between the plasma and the neutral gas. Introducing the effect of the friction into the induction equation yields the evolution equation (4.216) for the magnetic field, showing how the friction causes diffusion of the magnetic field.

4.7.1 Velocity-averaged scattering cross section

Suppose we now have a partially ionised medium, which can be seen as a mixture of neutral particles and plasma. As we have discussed before, the magnetic field can then be thought of being "frozen into" the plasma. Collisions between the plasma and neutral particles then create a friction force between the plasma and the neutral fluid, which causes the magnetic field to diffuse with respect to the plasma even if the plasma's conductivity is ideal. This diffusion process is called "ambipolar".

In order to work out this friction force, we first need a cross section σ for the collisions, or, more conveniently, the velocity-averaged cross section $\langle \sigma v \rangle$. We adopt two limiting cases for it, one for small and one for high relative velocities

$$v = \left| \vec{v}_\mathrm{i} - \vec{v}_\mathrm{n} \right| \tag{4.189}$$

during the interaction of the collision partners, i.e. the ions ("i") and the neutral gas particles ("n").

If v is very large, we may approximate the cross section by its geometrical value. If r_i and r_n are the effective radii of the ions and the neutral particles, respectively, we can then replace the cross section by a disk whose radius is the sum of the two radii,

$$\sigma = \pi(r_\mathrm{i} + r_\mathrm{n})^2 , \tag{4.190}$$

implying the velocity-averaged cross section

$$\langle\sigma v\rangle = \langle v\rangle\sigma = \langle v\rangle\pi(r_\mathrm{i} + r_\mathrm{n})^2 . \tag{4.191}$$

?

What would the cross section for the Coulomb scattering of an electron with an ion be, or between two ions?

If v is sufficiently small, the ion's charge can polarise the neutral particle and thereby enlarge the interaction cross section due to the electromagnetic interaction. While the electric field of an ion with charge Ze is the Coulomb field

$$\vec{E}_\mathrm{i} = \frac{Ze}{r^2}\hat{e}_r , \tag{4.192}$$

it appears reasonable to assume that the electric field of the polarised neutral particle is the dipole field

$$\vec{E}_\mathrm{n} = \frac{3(\vec{p}\cdot\hat{e}_r)\hat{e}_r - \vec{p}}{r^3} = -\vec{\nabla}\left(\frac{\vec{p}\cdot\vec{r}}{r^3}\right) \tag{4.193}$$

of the polarised dipole moment $\vec{p}$. We assume that the dipole moment responds linearly to the ion's electric field,

$$\vec{p} = \alpha\vec{E}_\mathrm{i} = \frac{Z\alpha e}{r^2}\hat{e}_r , \tag{4.194}$$

with a parameter α quantifying the polarisability of the neutral particles.

This induced dipole field of the neutral particle exerts the force

$$\vec{F} = Ze\vec{E}_\mathrm{n} = -Ze\vec{\nabla}\left(\frac{\vec{p}\cdot\vec{r}}{r^3}\right) = -Z^2e^2\alpha\vec{\nabla}\left(\frac{1}{r^4}\right) \tag{4.195}$$

on the ion, whose potential is evidently

$$V = \frac{Z^2e^2\alpha}{r^4} . \tag{4.196}$$

Since this is a central potential, Noether's theorems imply that motion under its influence conserves angular momentum. We can then characterise the motion of the ion past the neutral particle by the minimum separation r_0, the impact parameter b and the velocity v_∞ at infinity. Angular-momentum conservation requires

$$\mu v_\infty b = \mu r_0 v_0 , \tag{4.197}$$

where the reduced mass

$$\mu \equiv \frac{m_\mathrm{i}m_\mathrm{n}}{m_\mathrm{i} + m_\mathrm{n}} \tag{4.198}$$

occurs because we are treating a two-body problem. Energy conservation further demands

$$\frac{\mu}{2}v_\infty^2 = \frac{\mu}{2}v_0^2 - \frac{\alpha Z^2e^2}{r_0^4} . \tag{4.199}$$

Eliminating now the velocity v_0 at closest approach by angular-momentum conservation (4.197), we obtain a quadratic equation for the squared minimum separation r_0^2,

$$r_0^4 - b^2 r_0^2 + \frac{\alpha Z^2 e^2}{\mu v_\infty^2} = 0 , \tag{4.200}$$

which has the two solutions

$$r_{0,\pm}^2 = \frac{b^2}{2} \pm \sqrt{\frac{b^4}{4} - \frac{\alpha Z^2 e^2}{\mu v_\infty^2}} . \tag{4.201}$$

Both roots are mathematically possible, but only $r_{0,+}^2$ is physically relevant because in the limiting case of vanishing coupling α, the minimum radius must equal the impact parameter, $r_0 = b$, since the ion is then not scattered at all. The minimum separation r_0 will itself be smallest if the root in (4.201) vanishes and the impact parameter satisfies

$$b_0 = \left(\frac{4\alpha Z^2 e^2}{\mu v_\infty^2} \right)^{1/4} . \tag{4.202}$$

Since the force between the ion and the neutral particle decreases very steeply with increasing r, by far the strongest effect occurs for close encounters. Thus, we estimate the cross section as

$$\sigma = \pi b_0^2 = \frac{2\pi Z e}{v_\infty} \sqrt{\frac{\alpha}{\mu}} . \tag{4.203}$$

Obviously, the velocity-averaged cross section $\langle \sigma v_\infty \rangle$ is independent of the asymptotic velocity v_∞, and we find

$$\langle \sigma v_\infty \rangle = 2\pi Z e \sqrt{\frac{\alpha}{\mu}} . \tag{4.204}$$

4.7.2 Friction force and diffusion coefficient

A single collision between an ion and a neutral particle transfers the momentum

$$\left| \Delta \vec{p} \right| = \mu \left| \vec{v}_\mathrm{i} - \vec{v}_\mathrm{n} \right| \tag{4.205}$$

between the two. Since the scattering rate per volume is $n_\mathrm{i} n_\mathrm{n} \langle \sigma v_\infty \rangle$, the momentum transfer per unit time and volume is

$$\vec{f}_\mathrm{friction} = n_\mathrm{i} n_\mathrm{n} \langle \sigma v_\infty \rangle \mu \left(\vec{v}_\mathrm{i} - \vec{v}_\mathrm{n} \right) , \tag{4.206}$$

which corresponds to the spatial density of a friction force. With

$$\rho_\mathrm{i} \rho_\mathrm{n} = n_\mathrm{i} m_\mathrm{i} n_\mathrm{n} m_\mathrm{n} = (m_\mathrm{i} + m_\mathrm{n}) n_\mathrm{i} n_\mathrm{n} \mu , \tag{4.207}$$

this can be cast into the form

$$\boxed{\vec{f}_\mathrm{friction} = \gamma \rho_\mathrm{i} \rho_\mathrm{n} \left(\vec{v}_\mathrm{i} - \vec{v}_\mathrm{n} \right)} , \tag{4.208}$$

where the friction coefficient

$$\gamma \equiv \frac{\langle \sigma v_\infty \rangle}{m_\mathrm{i} + m_\mathrm{n}} \tag{4.209}$$

appears. In the two limiting cases discussed above, those of very small or very large relative velocities, we find

$$\gamma = \frac{\pi(r_\mathrm{i} + r_\mathrm{n})^2}{m_\mathrm{i} + m_\mathrm{n}} \left|\vec{v}_\mathrm{i} - \vec{v}_\mathrm{n}\right| \quad \text{or}$$
$$\gamma = \frac{2\pi Ze}{m_\mathrm{i} + m_\mathrm{n}} \sqrt{\frac{\alpha}{\mu}} \ . \tag{4.210}$$

A stationary situation can be established if this friction force between the ions and the neutral particles is balanced by the Lorentz force on the net electric current density caused by the motion of the plasma particles with the magnetic field. Since the Lorentz force density is

$$\vec{f}_\mathrm{L} = \frac{\vec{j} \times \vec{B}}{c} = \frac{1}{4\pi}\left(\vec{\nabla} \times \vec{B}\right) \times \vec{B} \ , \tag{4.211}$$

where Ampère's law was used in the second step, we find the relation

$$\vec{v}_\mathrm{d} \equiv \vec{v}_\mathrm{i} - \vec{v}_\mathrm{n} = \frac{\left(\vec{\nabla} \times \vec{B}\right) \times \vec{B}}{4\pi\gamma\rho_\mathrm{i}\rho_\mathrm{n}} \tag{4.212}$$

between the drift velocity of the ions relative to the neutral particles by equating the Lorentz force density (4.211) to the friction force density (4.208). In a magnetic field with characteristic length scale L, the drift velocity thus has the order of magnitude

$$v_\mathrm{d} \approx \frac{B^2}{4\pi\gamma\rho_\mathrm{i}\rho_\mathrm{n}L} \ . \tag{4.213}$$

A magnetic field which can be assumed to be "frozen" into the flow of an ideally conducting plasma must satisfy the induction equation

$$\frac{\partial \vec{B}}{\partial t} + \vec{\nabla} \times \left(\vec{B} \times \vec{v}_\mathrm{i}\right) = 0 \tag{4.214}$$

without the diffusion term arising from a finite conductivity. In order to calculate the diffusion of the magnetic field relative to the neutral particles, we transform into the rest frame of the neutral flow, where $\vec{v}_\mathrm{n} = 0$, allowing us to replace $\vec{v}_\mathrm{i}$ by the drift velocity $\vec{v}_\mathrm{d}$. This leads to the remarkable equation

$$\frac{\partial \vec{B}}{\partial t} + \vec{\nabla} \times \left(\vec{B} \times \frac{\left(\vec{\nabla} \times \vec{B}\right) \times \vec{B}}{4\pi\gamma\rho_\mathrm{i}\rho_\mathrm{n}}\right) = 0 \tag{4.215}$$

for the magnetic field in the rest-frame of the neutral gas. Using $\vec{\nabla} \cdot \vec{B} = 0$ and introducing the magnetic pressure $P_B = \vec{B}^2/(8\pi)$, it can be brought into the form

$$\frac{\partial \vec{B}}{\partial t} + \frac{1}{\gamma\rho_\mathrm{i}\rho_\mathrm{n}}\left[\vec{\nabla}\left(\vec{B} \cdot \vec{\nabla} P_B\right) - \vec{\nabla}^2\left(P_B\vec{B}\right)\right] = 0 \ . \tag{4.216}$$

The second term has the order of magnitude DB, where D corresponds to a diffusion coefficient

$$D \approx \frac{P_B}{\gamma\rho_\mathrm{i}\rho_\mathrm{n}} \ , \tag{4.217}$$

which approximately equals the estimate (4.213) for the drift velocity $\vec{v}_\mathrm{d}$, times the length scale L of the magnetic field.

Problems

1. Verify that (4.216) follows from (4.215).

4.8 Waves in magnetised cold plasmas

This section deals with electromagnetic waves in cold plasmas. Since the magnetic field imprints a preferred direction into the plasma which charges are gyrating around, the dielectric tensor is most conveniently split into contributions parallel, perpendicular and helical to the magnetic field, shown in (4.242). The dispersion relations (4.264) for electromagnetic waves are found to depend on the angle between their propagation direction and the magnetic field. We illustrate in (4.276) that waves polarised along their propagation direction are generally damped. Waves polarised transverse to their propagation direction are found to be eigenstates of the dielectric tensor only if they propagate along the magnetic field, in which case Faraday rotation occurs which is quantified by the rotation measure (4.284).

4.8.1 The dielectric tensor

We now proceed to study the propagation of electromagnetic waves in a magnetised plasma in which random particle motion is negligible, whose temperature is thus low, and which can in this sense be considered as cold. The equation of motion of an electron in such a plasma with embedded magnetic field $\vec{B}_0$ is then exclusively determined by the Lorentz force.

In a magnetised plasma irradiated by electromagnetic radiation, the Lorentz force is caused by the magnetic field embedded into the plasma, now called $\vec{B}_0$ to avoid confusion, together with the electric and magnetic fields, $\vec{E}$ and $\vec{B}$, of the incoming electromagnetic wave. The fields $\vec{E}$ and $\vec{B}$ of the electromagnetic wave are of similar magnitude. For non-relativistic motion of the plasma electrons, the magnetic part of the Lorentz force contributed by the electromagnetic wave can thus be neglected. Therefore, the Lorentz force of the combined fields has the electric part $-e\vec{E}$ of the electromagnetic wave and the magnetic part $-e\vec{\beta} \times \vec{B}_0$ of the magnetic field embedded into the plasma. The equation of motion for a plasma electron is then

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\frac{e}{m}\left(\vec{E} + \vec{\beta} \times \vec{B}_0\right) . \tag{4.218}$$

We now assume that the amplitude of the electric field $\vec{E}$ depends harmonically on time, $\vec{E}(t, \vec{x}) = \vec{E}(\vec{x})\mathrm{e}^{-\mathrm{i}\omega t}$. Likewise, we expand the velocity into an position-dependent amplitude and a harmonic time dependence, $\vec{v}(t, \vec{x}) = \vec{v}(\vec{x})\mathrm{e}^{-\mathrm{i}\omega t}$. Then, each of these monochromatic velocity modes is determined by

$$-\mathrm{i}\omega\vec{v} = -\frac{e}{m}\left(\vec{E} + \vec{\beta} \times \vec{B}_0\right) \tag{4.219}$$

due to the equation of motion (4.218). For the following calculations, it will be enormously helpful to introduce the matrix $\hat{\mathcal{B}}$ with the components

$$\hat{\mathcal{B}}_{ij} := \varepsilon_{ijk}\hat{b}^k , \tag{4.220}$$

where $\hat{b}$ is the unit vector in the direction of the magnetic field, $\vec{B}_0 = B_0\hat{b}$. This allows us to cast the linear system of equations (4.219) into the matrix form

$$\left(\mathrm{i}\omega\mathbb{1}_3 - \frac{eB_0}{mc}\hat{\mathcal{B}}\right)\vec{v} := M\vec{v} = \frac{e}{m}\vec{E}\ , \tag{4.221}$$

where we have separated the amplitude B_0 of the magnetic field from its unit direction vector $\hat{b}$. Identifying now the non-relativistic Larmor frequency

$$\omega_\mathrm{L} \equiv \frac{eB_0}{mc}\ , \tag{4.222}$$

see (2.60), we can abbreviate the matrix M implicitly defined in (4.221) by

$$M = \mathrm{i}\omega\mathbb{1}_3 - \omega_\mathrm{L}\hat{\mathcal{B}} = \mathrm{i}\omega\left(\mathbb{1}_3 + \mathrm{i}w_\mathrm{L}\hat{\mathcal{B}}\right)\ , \tag{4.223}$$

where $w_\mathrm{L} = \omega_\mathrm{L}/\omega$ is the Larmor frequency divided by the frequency of the incoming electromagnetic radiation. As we move on, the identities

$$\det\left(\mathbb{1}_3 + a\hat{\mathcal{B}}\right) = 1 + a^2\ , \quad \hat{\mathcal{B}}^2 = \hat{b}\otimes\hat{b} - \mathbb{1}_3 \quad \text{and} \quad \left(\hat{b}\otimes\hat{b}\right)\hat{\mathcal{B}} = 0 \tag{4.224}$$

will turn out to be most convenient. The first of these immediately gives

$$\det M = -\mathrm{i}\omega^3\left(1 - w_\mathrm{L}^2\right) = -\mathrm{i}\omega\left(\omega^2 - \omega_\mathrm{L}^2\right)\ , \tag{4.225}$$

while the second and the third will greatly simplify inverting the matrix M.

?

Verify the identities (4.224) by direct calculation.

The inverse of M should be a linear combination of the three matrices $\mathbb{1}_3$, $\hat{b}\otimes\hat{b}$, and $\hat{\mathcal{B}}$ we have available here. We thus try the ansatz

$$M^{-1} = A\mathbb{1}_3 + B\hat{b}\otimes\hat{b} + C\hat{\mathcal{B}}\ , \tag{4.226}$$

which must satisfy

$$\left(A\mathbb{1}_3 + B\hat{b}\otimes\hat{b} + C\hat{\mathcal{B}}\right)\left(\mathrm{i}\omega\mathbb{1}_3 + \omega_\mathrm{L}\hat{\mathcal{B}}\right) = \mathbb{1}_3\ . \tag{4.227}$$

Multiplying out the sums in the parentheses, using the identities (4.224) and grouping terms takes us immediately to

$$A\mathrm{i}\omega - C\omega_\mathrm{L} = 1\ , \quad B\mathrm{i}\omega + C\omega_\mathrm{L} = 0 \quad \text{and} \quad C\mathrm{i}\omega + A\omega_\mathrm{L} = 0\ . \tag{4.228}$$

The sum of the first and the second of these equations implies

$$\mathrm{i}\omega(A + B) = 1\ , \tag{4.229}$$

while multiplying the first with $\mathrm{i}\omega_\mathrm{L}$ and the third with ω and adding these two results in

$$C = \frac{w_\mathrm{L}}{\omega\left(1 - w_\mathrm{L}^2\right)}\ . \tag{4.230}$$

?

Determine the coefficients (A, B, C) from (4.226) in your own, independent calculation. Confirm that the matrix M^{-1} found in (4.232) indeed inverts the matrix M defined in (4.221).

Substituting backwards and identifying $1 - w_\mathrm{L}^2 = \mathrm{i}\omega^{-3}\det M$ from (4.225) gives

$$A = -\frac{\omega^2}{\det M}\ , \quad B = Aw_\mathrm{L}^2 \quad \text{and} \quad C = -\mathrm{i}Aw_\mathrm{L}\ . \tag{4.231}$$

Thus, the matrix M is inverted by

$$M^{-1} = \frac{\omega^2}{\det M}\left(-\mathbb{1}_3 + w_\mathrm{L}^2\hat{b}\otimes\hat{b} - \mathrm{i}w_\mathrm{L}\hat{\mathcal{B}}\right)\ . \tag{4.232}$$

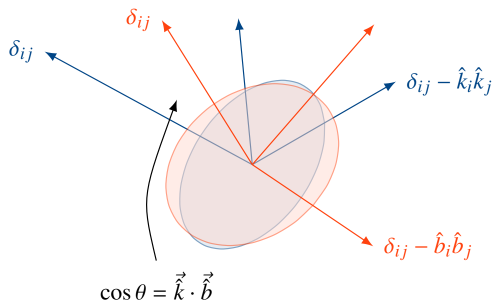

Figure 4.7 Electromagnetic waves in a magnetised plasma experience two preferred directions which are usually not aligned.

With this result, we can invert (4.221) to obtain the velocity

$$\vec{v} = \frac{e}{m} M^{-1}\vec{E} = \frac{e}{\mathrm{i}m\omega\left(1 - w_\mathrm{L}^2\right)}\left(-\mathbb{1}_3 + w_\mathrm{L}^2 \hat{b} \otimes \hat{b} - \mathrm{i}w_\mathrm{L}\hat{\mathcal{B}}\right) . \tag{4.233}$$

This electron velocity gives us the polarised current density $\vec{j}_\mathrm{pol}$, from which we can find the polarisation and the dielectricity tensor.

As usual, the polarisation is given by $4\pi\vec{P} = \vec{D} - \vec{E}$, and its derivative with respect to time is the polarised current density

$$\vec{j}_\mathrm{pol} = -en_\mathrm{e}\vec{v} = \frac{\partial\vec{P}}{\partial t} . \tag{4.234}$$

Extending the assumed harmonic time dependence of the incoming electromagnetic radiation to the polarisation, we adopt $\vec{P}(t, \vec{x}) = \vec{P}(\vec{x})\mathrm{e}^{-\mathrm{i}\omega t}$. Then, the previous equation shows that the electromagnetic fields and the velocity must be related by

$$-\mathrm{i}\omega\vec{P} = -\frac{\mathrm{i}\omega}{4\pi}\left(\vec{D} - \vec{E}\right) = -en_\mathrm{e}\vec{v} . \tag{4.235}$$

Solving the last equation in this chain for the dielectric displacement $\vec{D}$, and substituting the velocity vector from (4.233), we find

$$\vec{D} = \frac{4\pi en_\mathrm{e}}{\mathrm{i}\omega}\vec{v} + \vec{E} = \vec{E} - \frac{w_\mathrm{p}^2}{1 - w_\mathrm{L}^2}\left(\mathbb{1}_3 - w_\mathrm{L}^2 \hat{b} \otimes \hat{b} - \mathrm{i}w_\mathrm{L}\hat{\mathcal{B}}\right)\vec{E} , \tag{4.236}$$

where we have identified the plasma frequency

$$\omega_\mathrm{p} = \sqrt{\frac{4\pi e^2 n_\mathrm{e}}{m}} \tag{4.237}$$

and introduced its dimension-less form $w_\mathrm{p} = \omega_\mathrm{p}/\omega$. Recalling that the dielectric tensor is defined by the linear relation $\vec{D} = \varepsilon\vec{E}$, we can directly read it off (4.236) and obtain

$$\varepsilon = \left(1 - \frac{w_\mathrm{p}^2}{1 - w_\mathrm{L}^2}\right)\mathbb{1}_3 + \frac{w_\mathrm{p}^2 w_\mathrm{L}^2}{1 - w_\mathrm{L}^2}\,\hat{b} \otimes \hat{b} + \frac{\mathrm{i}w_\mathrm{p}^2 w_\mathrm{L}}{1 - w_\mathrm{L}^2}\,\hat{\mathcal{B}} . \tag{4.238}$$

This result for the dielectric tensor can be further decomposed into components parallel and perpendicular to the magnetic field, and an additional, antisymmetric contribution. To this end, we define the parallel and perpendicular projection operators (Figure 4.7),

$$\pi_{\parallel} := \hat{b} \otimes \hat{b}\ , \quad \pi_{\perp} := \mathbb{1}_3 - \hat{b} \otimes \hat{b}\ , \tag{4.239}$$

and contract ε with them, finding

$$\varepsilon_{\parallel} = \mathrm{Tr}\,(\pi_{\parallel}\varepsilon) = 1 - w_{\mathrm{p}}^2\ , \quad \varepsilon_{\perp} = \frac{1}{2}\,\mathrm{Tr}\,(\pi_{\perp}\varepsilon) = 1 - \frac{w_{\mathrm{p}}^2}{1 - w_{\mathrm{L}}^2}\ . \tag{4.240}$$

?

Can you confirm the expressions (4.240) for the longitudinal and the transverse dielectricity?

Finally abbreviating the antisymmetric amplitude by

$$g = \frac{w_{\mathrm{p}}^2 w_{\mathrm{L}}}{1 - w_{\mathrm{L}}^2}\ , \tag{4.241}$$

we can bring the dielectricity tensor into the compact form

$$\varepsilon = \left(1 - w_{\mathrm{p}}^2\right)\pi_{\parallel} + \left(1 - \frac{w_{\mathrm{p}}^2}{1 - w_{\mathrm{L}}^2}\right)\pi_{\perp} + \mathrm{i}\frac{w_{\mathrm{p}}^2 w_{\mathrm{L}}}{1 - w_{\mathrm{L}}^2}\,\hat{\mathcal{B}}\ . \tag{4.242}$$

4.8.2 Contribution by ions

If ions need to be taken into consideration, the parallel, perpendicular and antisymmetric dielectricity components change according to

$$\begin{aligned} \varepsilon_{\perp} - 1 &\to (\varepsilon_{\perp} - 1)_{\mathrm{e}} + (\varepsilon_{\perp} - 1)_{\mathrm{i}}\ , \\ \varepsilon_{\parallel} - 1 &\to (\varepsilon_{\parallel} - 1)_{\mathrm{e}} + (\varepsilon_{\parallel} - 1)_{\mathrm{i}}\ , \\ g &\to g_{\mathrm{e}} + g_{\mathrm{i}}\ , \end{aligned} \tag{4.243}$$

where the plasma and the Larmor frequencies of the electrons and the ions have to be distinguished. The Larmor frequency of ions with charge Ze is

$$\omega_{\mathrm{L,i}} = \frac{ZeB_0}{m_{\mathrm{i}}c} = f\omega_{\mathrm{L,e}} \ll \omega_{\mathrm{L,e}} \quad \text{with} \quad f \equiv \frac{Zm_{\mathrm{e}}}{m_{\mathrm{i}}}\ , \tag{4.244}$$

much smaller than the Larmor frequency of the electrons. The squared plasma frequency of the ions is

$$\omega_{\mathrm{p,i}}^2 = \frac{4\pi Z^2 e^2 n_{\mathrm{i}}}{m_{\mathrm{i}}} = \frac{4\pi Z e^2 n_{\mathrm{e}}}{m_{\mathrm{i}}} = f\omega_{\mathrm{p,e}}^2\ , \tag{4.245}$$

where we have used in the second step that the plasma is supposed to be neutral, $Zen_{\mathrm{i}} = en_{\mathrm{e}}$. Therefore, the ratio between the plasma frequencies of the ions and the electrons is

$$\frac{\omega_{\mathrm{pi}}}{\omega_{\mathrm{pe}}} = f^{1/2} \ll 1\ , \tag{4.246}$$

also much less than unity.

The contribution of the ions to the longitudinal dielectricity $\varepsilon_{\|}$ thus turns out to be negligible. However, their contribution to the transverse dielectricity $\varepsilon_{\perp}$ and the antisymmetric dielectricity g is not necessarily small. The ratio

$$\left|\left(\frac{w_{\mathrm{p,i}}^2}{1-w_{\mathrm{L,i}}^2}\right)\left(\frac{w_{\mathrm{p,e}}^2}{1-w_{\mathrm{L,e}}^2}\right)^{-1}\right| = f\left|\frac{1-w_{\mathrm{L,e}}^2}{1-f^2w_{\mathrm{L,e}}^2}\right| \tag{4.247}$$

is of order unity if

$$\left|\frac{1-w_{\mathrm{L,e}}^2}{1-f^2w_{\mathrm{L,e}}^2}\right| \approx \frac{1}{f} \tag{4.248}$$

holds. Searching for a solution of this approximate equation, it turns out that we need to choose the negative branch of the modulus on the left-hand side because otherwise w_{L}^2 turned out negative. Then, the approximate equation (4.248) demands

$$-f\left(1-w_{\mathrm{L,e}}^2\right) \approx 1-f^2w_{\mathrm{L,e}}^2\;, \tag{4.249}$$

showing that ions can contribute substantially to the transverse dielectricity if the Larmor frequency of the electrons is much larger than the frequency ω of the electromagnetic radiation,

$$w_{\mathrm{L,e}} = \frac{\omega_{\mathrm{L,e}}}{\omega} \lesssim \frac{1}{\sqrt{f}}\;. \tag{4.250}$$

In a fully analogous way, we can see that the contributions of ions and electrons to g are comparable if the radiation frequency ω is suitably small compared to $\omega_{\mathrm{L,e}}$,

$$w_{\mathrm{L}}^2 \lesssim \frac{1}{2f^2}\;. \tag{4.251}$$

Thus, radiation with sufficiently low frequency, $\omega \lesssim \sqrt{f}\omega_{\mathrm{L,e}}$, will feel the ion contribution to the transverse dielectricity, and it will feel the ion contribution to the antisymmetric dielectricity if $\omega \lesssim \sqrt{2}f\omega_{\mathrm{L,e}}$. For a pure hydrogen plasma, $f \approx 5.6 \cdot 10^{-4}$. Ions then become important for the transverse dielectricity for frequencies $\omega \lesssim 0.02\omega_{\mathrm{L,e}}$, and for the antisymmetric dielectricity g if $\omega \lesssim 8 \cdot 10^{-4}\omega_{\mathrm{L,e}}$.

4.8.3 Dispersion relations in a cold, magnetised plasma

With the dielectric tensor (4.239) in presence of a magnetic field, we return to the general dispersion relation (4.49)

$$\det\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k} - \frac{\omega^2}{k^2c^2}\hat{\varepsilon}\right) = 0\;. \tag{4.252}$$

Through the dielectric tensor, this dispersion relation contains the projectors $\pi_{\|}$ and $\pi_{\perp}$ parallel and perpendicular to the magnetic field, while the first two terms together are the projector

$$\tilde{\pi}_{\perp} = \mathbb{1}_3 - \hat{k}\otimes\hat{k} \tag{4.253}$$

perpendicular to the propagation direction of the incoming electromagnetic field, which we now mark with a tilde to distinguish it from the projectors relative to the magnetic field. It is advantageous to express $\tilde{\pi}_\perp$ in terms of the projectors $\pi_\parallel$ and $\pi_\perp$. To this end, we write

$$\tilde{\pi}_\perp = A\pi_\parallel + B\pi_\perp \ , \tag{4.254}$$

apply $\pi_\parallel$ and $\pi_\perp$ from the left and take the trace to find

$$A = \mathrm{Tr}\,(\pi_\parallel\tilde{\pi}_\perp) \quad \text{and} \quad B = \frac{1}{2}\,\mathrm{Tr}\,(\pi_\perp\tilde{\pi}_\perp) \ . \tag{4.255}$$

Caution Notice (and convince yourself) that $\mathrm{Tr}(\hat{b}\otimes\hat{b}) = \hat{b}^2 = 1$ and

$$\mathrm{Tr}\left[\left(\hat{b}\otimes\hat{b}\right)\left(\hat{k}\otimes\hat{k}\right)\right] = \left(\hat{b}\cdot\hat{k}\right)^2 = \cos^2\theta \ .$$

◄

The remaining traces are easily worked out. To do so, we introduce the angle θ between the direction $\hat{k}$ of the wave propagation and $\hat{b}$ of the magnetic field by $\cos\theta = \hat{k}\cdot\hat{b}$.

This gives

$$A = \mathrm{Tr}\,(\pi_\parallel\tilde{\pi}_\perp) = \mathrm{Tr}\left(\hat{b}\otimes\hat{b}\right)\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k}\right) = 1 - \cos^2\theta = \sin^2\theta \tag{4.256}$$

and

$$B = \frac{1}{2}\,\mathrm{Tr}\,(\pi_\perp\tilde{\pi}_\perp) = \frac{1}{2}\,\mathrm{Tr}\left(\mathbb{1}_3 - \hat{b}\otimes\hat{b}\right)\left(\mathbb{1}_3 - \hat{k}\otimes\hat{k}\right) = \frac{1+\cos^2\theta}{2} \ , \tag{4.257}$$

allowing us to write the projector $\tilde{\pi}_\perp$ perpendicular to the wave vector as the linear combination

$$\tilde{\pi}_\perp = \sin^2\theta\,\pi_\parallel + \frac{1+\cos^2\theta}{2}\,\pi_\perp \tag{4.258}$$

of the projectors $\pi_\perp$ and $\pi_\parallel$ relative to the magnetic field. In much the same way, the parallel projector $\tilde{\pi}_\parallel$ relative to the wave vector can be expanded as

$$\tilde{\pi}_\parallel = \cos^2\theta\,\pi_\parallel + \frac{\sin^2\theta}{2}\,\pi_\perp \ . \tag{4.259}$$

?

Perform the calculations yourself that lead to the representations (4.258) and (4.259) of the perpendicular and parallel projectors with respect to $\hat{k}$.

Introducing (4.258) into (4.252) and abbreviating further $\omega^2/(k^2c^2) =: w^2$ enables us to write the dispersion relation as

$$\det\left\{\left(\sin^2\theta - w^2\varepsilon_\parallel\right)\pi_\parallel + \left[\frac{1}{2}\left(1+\cos^2\theta\right) - w^2\varepsilon_\perp\right]\pi_\perp - \mathrm{i}w^2g\hat{\mathcal{B}}\right\} = 0 \ . \tag{4.260}$$

Since the determinant is invariant under orthogonal transformations, we can rotate the coordinate frame such that $\hat{b}$ points into the $\hat{e}_z$ direction such that $\hat{b}^i = \delta^i_3$. Under this choice, which does not affect generality, the dispersion relation (4.260) simplifies to

$$\det\begin{pmatrix} \frac{1}{2}\left(1+\cos^2\theta\right) - w^2\varepsilon_\perp & -\mathrm{i}w^2g & 0 \\ \mathrm{i}w^2g & \frac{1}{2}\left(1+\cos^2\theta\right) - w^2\varepsilon_\perp & 0 \\ 0 & 0 & \sin^2\theta - w^2\varepsilon_\parallel \end{pmatrix} = 0 \ . \tag{4.261}$$

The solutions can directly be read off this expression now. They are

$$\sin^2\theta = w^2\varepsilon_\parallel \quad \text{and} \quad \frac{1+\cos^2\theta}{2} = w^2\left(\varepsilon_\perp \pm g\right) \ , \tag{4.262}$$

showing that the propagation of electromagnetic waves through a magnetised plasma depends on the propagation direction of the waves relative to the magnetic field. Inserting the expressions (4.240) and (4.241) for the parallel, transverse, and antisymmetric dielectricities, the dispersion relations (4.262) can finally be cast into the forms

$$w^2\left(1-w_\mathrm{p}^2\right)=\sin^2\theta \quad\text{and}\quad w^2\left(1-\frac{w_\mathrm{p}^2}{1\pm w_\mathrm{L}}\right)=\frac{1+\cos^2\theta}{2} \tag{4.263}$$

or, when written in terms of the frequencies,

$$\omega^2-\omega_\mathrm{p}^2=k^2c^2\sin^2\theta \quad\text{and}\quad \omega^2\left[1-\frac{\omega_\mathrm{p}^2}{\omega(\omega\pm\omega_\mathrm{L})}\right]=\frac{1+\cos^2\theta}{2}\,k^2c^2\;. \tag{4.264}$$

The first dispersion relation (4.264) is a second-order polynomial with one real root. It applies to waves with an electric field vector polarised parallel to the magnetic field. The second dispersion relation is a fourth-order polynomial with four real roots which belong to two branches split by the Larmor frequency of gyration in the magnetic field (Figure 4.8).

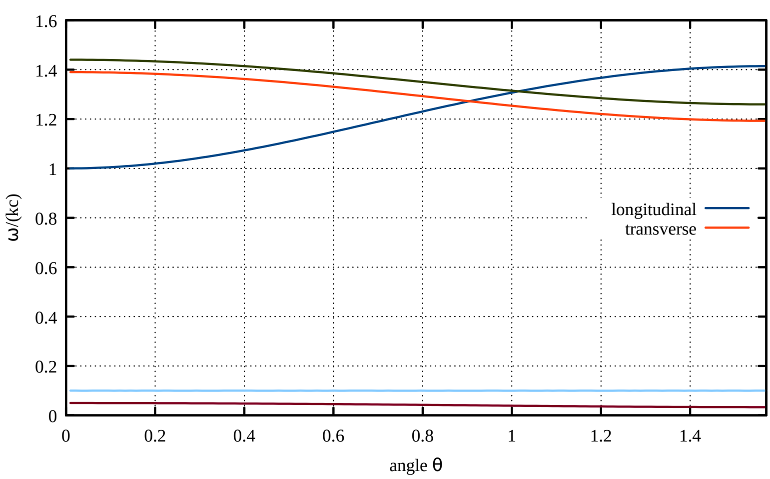

Figure 4.8 Dispersion relations for electromagnetic waves propagating through a magnetised, cold plasma.

4.8.4 Longitudinal and transverse waves

Having derived the dispersion relations for waves polarised relative to the magnetic field, it is now very interesting and instructive to repeat this derivation for waves polarised longitudinally or transversally to their propagation direction. For this purpose, we need to return to the original propagation equation

$$\left(\tilde{\pi}_\perp-w^2\varepsilon\right)\hat{\vec{E}}=0\;, \tag{4.265}$$

but now expand the projectors $\pi_\|$ and $\pi_\perp$ relative to the magnetic field in terms of the projectors $\tilde{\pi}_\|$ and $\tilde{\pi}_\perp$ parallel and perpendicular to the propagation direction

$\hat{k}$. By a procedure entirely analogous to the reverse expansion of $\tilde{\pi}_\perp$ performed above, we find expressions (4.258) and (4.259) with the projectors $\tilde{\pi}$ and π interchanged,

$$\pi_\perp = \sin^2\theta\,\tilde{\pi}_\| + \frac{1+\cos^2\theta}{2}\,\tilde{\pi}_\perp \quad \text{and} \quad \pi_\| = \cos^2\theta\,\tilde{\pi}_\| + \frac{\sin^2\theta}{2}\,\tilde{\pi}_\perp \tag{4.266}$$

and accordingly decompose the propagation equation (4.265) as

$$\begin{aligned}\Bigg\{\left(1 - \frac{w^2}{2}\left[\varepsilon_\| \sin^2\theta + \varepsilon_\perp\left(1+\cos^2\theta\right)\right]\right)\tilde{\pi}_\perp \\ - w^2\left(\varepsilon_\|\cos^2\theta + \varepsilon_\perp\sin^2\theta\right)\tilde{\pi}_\| - \mathrm{i}w^2 g\,\hat{\mathcal{B}}\Bigg\}\hat{\vec{E}} = 0\,.\end{aligned} \tag{4.267}$$

We turn the coordinate system such that the wave propagates into the z direction, $\hat{k} = \hat{e}_z$, and that the unit vector $\hat{b}$ in field direction falls into the $x - z$ plane,

$$\hat{b} = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix}\,. \tag{4.268}$$

The propagation condition (4.267) can then be cast into the matrix form

$$\begin{pmatrix} A & -\mathrm{i}C\cos\theta & 0 \\ \mathrm{i}C\cos\theta & A & \mathrm{i}C\sin\theta \\ 0 & -\mathrm{i}C\sin\theta & -B \end{pmatrix}\begin{pmatrix} \hat{E}_x \\ \hat{E}_y \\ \hat{E}_z \end{pmatrix} = 0 \tag{4.269}$$

with the abbreviations

$$\begin{aligned} A &= 1 - \frac{w^2}{2}\left[\varepsilon_\|\sin^2\theta + \varepsilon_\perp\left(1+\cos^2\theta\right)\right]\,, \\ B &= w^2\left(\varepsilon_\|\cos^2\theta + \varepsilon_\perp\sin^2\theta\right) \quad \text{and} \\ C &= w^2 g\,. \end{aligned} \tag{4.270}$$

?

Verify that, for the orientation of the coordinate frame defined in the text, (4.267) turns into (4.269) with the coefficients (4.270).

Longitudinal waves. Consider now a longitudinally polarised wave, $\hat{E}_x = 0 = \hat{E}_y$. According to (4.270), its dispersion relation is $B - \mathrm{i}C\cos\theta = 0$ or

$$\varepsilon_\|\cos^2\theta + \varepsilon_\perp\sin^2\theta - \mathrm{i}g\sin\theta = 0\,. \tag{4.271}$$

An imaginary part appears if the wave does not propagate in or against the direction of the magnetic field, $\theta = 0$ or $\theta = \pi$. Then, the frequency is typically complex, and the longitudinally polarised wave is expected to be damped. Let us have a closer look into this. If $\sin\theta = 0$, the dispersion relation simplifies to $\varepsilon_\| = 0$ or $\omega = \omega_\mathrm{p}$ according to (4.240). Such waves can propagate if they have the plasma frequency. Suppose now $\sin\theta \neq 0$, but let the Larmor frequency be much smaller than the plasma frequency, allowing us to approximate

$$\varepsilon_\perp \approx 1 - w_\mathrm{p}^2 = \varepsilon_\|\,, \quad g \approx w_\mathrm{p}^2 w_\mathrm{L} \tag{4.272}$$

to linear order in w_L. We assume that the frequency ω can then be approximated as $\omega = \omega_\mathrm{p} + \delta\omega$ with a small and possibly complex correction $\delta\omega$. To first order in $\delta\omega$ and ω_L, the dielectricities are

$$\varepsilon_\perp \approx \varepsilon_\| = 1 - \frac{\omega_\mathrm{p}^2}{\omega^2} = 1 - \frac{1}{\left(1 + \delta\omega/\omega_\mathrm{p}\right)^2} \approx 2\frac{\delta\omega}{\omega_\mathrm{p}} \tag{4.273}$$

and

$$g = \frac{\omega_p^2 \omega_L}{\omega^3} = \frac{\omega_L}{\omega_p \left(1 + \delta\omega/\omega_p\right)^3} \approx \frac{\omega_L}{\omega_p} \left(1 - 3\frac{\delta\omega}{\omega_p}\right) . \tag{4.274}$$

With these expressions, the dispersion relation (4.271) turns into

$$2\frac{\delta\omega}{\omega_p} - \mathrm{i}\frac{\omega_L}{\omega_p}\left(1 - 3\frac{\delta\omega}{\omega_p}\right)\sin\theta = 0 , \tag{4.275}$$

which has the solution

$$\delta\omega \approx \frac{\mathrm{i}}{2}\omega_L \sin\theta \tag{4.276}$$

to first order in the Larmor frequency ω_L. This imaginary part damps the longitudinally polarised waves.

Transverse waves. Beginning instead with transversally polarised waves, $\hat{E}_z = 0$, we see immediately that field vectors in the x-y plane cannot be eigenvectors of the matrix in (4.269) unless $\sin\theta = 0$: An initially transversally polarised field vector $\hat{E}$ immediately acquires a longitudinal component $\hat{E}_z = \mathrm{i}\hat{E}_y C \sin\theta$. For transverse waves to remain transverse, let us therefore assume that the wave vector is aligned with the magnetic field, $\sin\theta = 0$. Then, the propagation condition (4.269) shrinks to

$$\begin{pmatrix} 1 - w^2\varepsilon_\perp & -\mathrm{i}w^2 g \\ \mathrm{i}w^2 g & 1 - w^2\varepsilon_\perp \end{pmatrix} \begin{pmatrix} \hat{E}^1 \\ \hat{E}^2 \end{pmatrix} = 0 . \tag{4.277}$$

Using the second dispersion relation from (4.262) with $\cos\theta = \pm 1$, we can eliminate

$$w^2 = \frac{1}{\varepsilon_\perp \pm g} \tag{4.278}$$

and bring the condition (4.277) into the simple form

$$\frac{g}{\varepsilon_\perp \pm g} \begin{pmatrix} \pm 1 & -\mathrm{i} \\ \mathrm{i} & \pm 1 \end{pmatrix} \begin{pmatrix} \hat{E}^1 \\ \hat{E}^2 \end{pmatrix} = 0 . \tag{4.279}$$

This equation immediately shows that the electric-field components must satisfy

$$\hat{E}^1 = \mp \mathrm{i}\hat{E}^2 \quad \text{or} \quad \left(\hat{E}^1\right)^2 + \left(\hat{E}^2\right)^2 = 1 , \tag{4.280}$$

which characterises circularly polarised light: The two remaining components of the electric field are determined such that the electric-field vector lies on a circle. Since the multiplication with the imaginary unit i corresponds to a rotation by $\pi/2$ in the plane transversal to the propagation direction of the wave, $E^1 = \mp \mathrm{i} E^2$ describe right- and left-circularly polarised light.

4.8.5 Faraday rotation

The preceding discussion of transversal waves propagating parallel to the magnetic field thus leads us to the conclusion that the two branches of the second dispersion relation from (4.263) describe the propagation of left- and right-circularly polarised waves which propagate differently because they obey different dispersion relations. The left-circular polarisation state propagates through the magnetised plasma in a different way than the right-circular polarisation state. Qualitatively, this is not surprising because the motion of the electrons in the magnetised plasma has a fixed sense of rotation: The Lorentz force (1.146) on negatively charged particles requires them to spiral counter-clockwise around the magnetic field lines, seen in the direction of the field lines themselves.

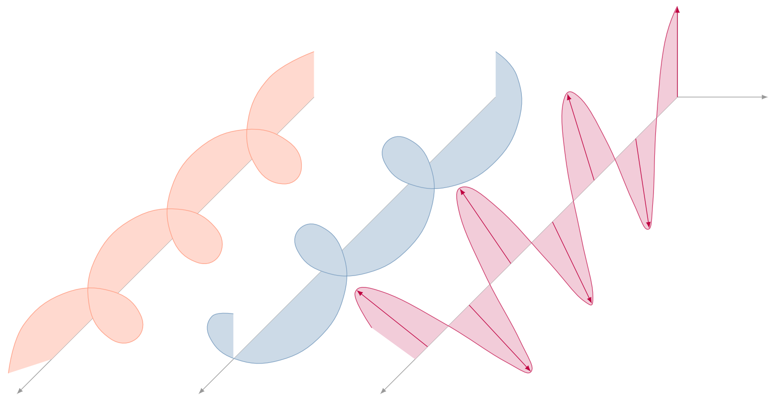

Figure 4.9 The superposition of a left- and a right-circularly polarised wave with a slight phase difference is a plane-polarised wave with a rotating polarisation direction.

Linearly polarised light can be decomposed into left- and right-circularly polarised modes of equal intensity and constant phase difference. If the two circularly-polarised modes now travel through a magnetised plasma at different phase velocities, their phase difference changes as they travel. The polarisation direction of linearly polarised light is then rotated. This effect is called *Faraday rotation* (Figure 4.9). In the high-frequency limit, when the frequency of the electromagnetic wave is much higher than both the Larmor and the plasma frequencies, $\omega \gg \omega_\mathrm{L}$ and $\omega \gg \omega_\mathrm{p}$, we can approximate the second dispersion relation from (4.264) for $\cos\theta = \pm 1$ by

$$k_\pm^2 = \frac{\omega^2}{c^2}\left[1 - \frac{\omega_\mathrm{p}^2}{\omega(\omega \pm \omega_\mathrm{L})}\right] \approx \frac{\omega^2}{c^2}\left[1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\left(1 \mp \frac{\omega_\mathrm{L}}{\omega}\right)\right] \tag{4.281}$$

in a first step for which only $\omega \gg \omega_\mathrm{L}$ is required. The condition $\omega \gg \omega_\mathrm{p}$ allows us to continue by taking a square root to first order Taylor approximation,

$$k_\pm \approx \frac{\omega}{c}\left[1 - \frac{\omega_\mathrm{p}^2}{2\omega^2}\left(1 \mp \frac{\omega_\mathrm{L}}{\omega}\right)\right] = \left(\frac{\omega}{c} - \frac{\omega_\mathrm{p}^2}{2\omega c}\right) \pm \frac{\omega_\mathrm{p}^2\omega_\mathrm{L}}{2\omega^2 c} \equiv k_0 \pm \Delta k \,. \tag{4.282}$$

The first term, k_0, corresponds to a wave vector in the unmagnetised plasma, while the second term, Δk, is responsible for the phase shift between the left- and right-circularly polarised states. This phase shift causes the direction of linear polarisation by an angle

$$\psi = \int \Delta k \mathrm{d}z = \int \frac{\omega_\mathrm{p}^2\omega_\mathrm{L}}{2\omega^2 c}\mathrm{d}z = \int \frac{4\pi e^2 n_\mathrm{e}}{m}\frac{eB}{mc}\frac{\mathrm{d}z}{2\omega^2 c} = \frac{2\pi e^3}{m^2c^2\omega^2}\int \mathrm{d}z\, n_\mathrm{e} B \,, \tag{4.283}$$

where we have inserted the explicit expressions for the plasma and Larmor frequencies. According to this result, in the high-frequency limit, the Faraday rotation is proportional to ω^{-2} or, equivalently, to the squared wave length λ^2. The expression

$$\int \mathrm{d}z\, n_\mathrm{e} B \equiv \mathrm{RM} \tag{4.284}$$

is called the *rotation measure*.

Faraday rotation is an important diagnostic for astrophysical magnetic fields. If a source of linearly polarised light, such as a radio source emitting synchrotron radiation, shines through a magnetised plasma in its foreground, the plane of linear polarisation rotates by different amounts at different frequencies. If the polarisation direction can be measured in two or more frequency bands, the rotation measure can be determined and thus a line-of-sight integral over the magnetic field strength parallel to the line-of-sight, weighted by the electron density. Assumptions then need to be made on the orientation of the magnetic field and on the electron density, under which the field strength can then be estimated.

Problems

1. Return to the dispersion relations (4.264) for electromagnetic waves in a magnetised plasma and consider their limit for very weak fields. Derive approximate dispersion relations for this case and discuss their physical meaning.

2. Since Faraday rotation is only sensitive to the line-of-sight component $B_\|$ of the magnetic field, it can only measure a net magnetic field remaining after cancellation of sections along the line-of-sight where the field is pointing towards and away from the observer.

   (a) What is the expectation value of the rotation measure created by a randomly magnetised, homogeneous medium?

   (b) What is the variance of the distribution of rotation measures obtained along many different lines-of-sight through the randomly magnetised medium if the energy density in the magnetic field is U_B?

   (c) If the magnetic field is not completely random, but has a correlation function $\xi_B(r)$ given by

   $$\xi_B(r)\delta_{ij} = \left\langle B_i\,(\vec x)\,B_j\,(\vec x + \vec r)\right\rangle\ , \tag{4.285}$$

   what correlation function of the rotation measure is observed?

   (d) Suppose a magnetised medium of thickness L can be modelled as composed of subvolumes with a characteristic linear dimension λ, carrying magnetic fields of identical strength B_0 but random orientation. How will the variance of the observed distribution of Faraday rotations depend on λ and L?

4.9 Hydromagnetic Waves

In this section, a linear perturbation analysis of the equations of ideal, inviscid magnetohydrodynamics is carried out, allowing us to identify different modes of hydromagnetic waves. The result of linearising the equations in the perturbations is the general dispersion relation (4.301), which can be specialised to identify the dispersion relation (4.304) for Alfvén waves and to infer the existence of fast and slow hydromagnetic waves with the sound speeds (4.309).

4.9.1 Linearised perturbation equations

Now we consider, in a way very similar to the treatment of sound waves in a neutral fluid, the propagation of waves in a magnetised plasma. For simplicity, we assume that dissipation and heat conduction are unimportant, $\zeta = \eta = \kappa = 0$, and that the conductivity be infinite, $\sigma^{-1} = 0$. Then, the combined equations of this ideal, inviscid specialisation of magnetohydrodynamics read

$$\vec\nabla \cdot \vec B = 0\ , \quad \frac{\partial \vec B}{\partial t} = \vec\nabla \times \left(\vec v \times \vec B\right)\ , \quad \frac{\partial \rho}{\partial t} + \vec\nabla \cdot (\rho \vec v) = 0\ ,$$
$$\frac{\partial \vec v}{\partial t} + \left(\vec v \cdot \vec\nabla\right)\vec v = -\frac{\vec\nabla P}{\rho} + \frac{1}{4\pi\rho}\left(\vec\nabla \times \vec B\right) \times \vec B\ : \tag{4.286}$$

Besides mass conservation and Maxwell's equation $\vec\nabla \cdot \vec B = 0$, the magnetic field must satisfy the induction equation and Euler's equation must contain the back-reaction of the magnetic field on the plasma flow. The energy conservation equation is not relevant for the following considerations. We now proceed as usual in a perturbation analysis. We begin by assuming that an equilibrium solution for the magnetic field and the plasma quantities exists,

$$\vec B_0\ , \quad \rho_0\ , \quad P_0\ , \quad \vec v_0 = 0\ , \tag{4.287}$$

which we indicate by the subscript 0. Setting the equilibrium velocity to zero is not a severe restriction because it means that we transform into a coordinate system comoving with the equilibrium plasma flow. This equilibrium solution is then perturbed by small amounts

$$\delta\vec B\ , \quad \delta\rho\ , \quad \delta P\ , \quad \delta\vec v \tag{4.288}$$

in all variables. In absence of dissipation, entropy has to be conserved along flow lines. We further assume isentropic flow, thus $s = \text{const.}$ everywhere in the flow.

Then, we proceed by linearising the ideal magnetohydrodynamic equations. For doing so, we insert the perturbed variables $\vec B_0 + \delta\vec B$, $\vec\rho_0 + \delta\rho$, $P_0 + \delta P$ and $\delta\vec v$ into the equations (4.286) and drop all terms of higher than first order in the perturbations. Moreover, we use the fact that the equilibrium quantities $\vec B_0$, ρ_0, P_0 and $\vec v_0$ are in fact solutions of the equations. In this way, we find from the first three equations (4.286)

$$\vec\nabla \cdot \delta\vec B = 0\ , \quad \frac{\partial \delta\vec B}{\partial t} = \vec\nabla \times \left(\delta\vec v \times \vec B_0\right)\ , \quad \frac{\partial \delta\rho}{\partial t} + \vec\nabla \cdot (\rho_0 \delta\vec v) = 0\ . \tag{4.289}$$

Suppose further that the fluctuations in the density, $\delta\rho$, are of much smaller scale than any scale on which the equilibrium density ρ_0 might change. This allows us to assume that the equilibrium density is locally constant, $\rho_0 = \text{const.}$, so that the continuity equation can be simplified to read

$$\frac{\partial \delta\rho}{\partial t} + \rho_0 \vec{\nabla} \cdot \delta\vec{v} = 0 \ . \tag{4.290}$$

Finally, also to first order in all perturbations, Euler's equation reads

$$\frac{\partial \delta\vec{v}}{\partial t} = -\frac{\vec{\nabla}\delta P}{\rho_0} + \frac{\left(\vec{\nabla} \times \delta\vec{B}\right) \times \vec{B}_0}{4\pi\rho_0} \ , \tag{4.291}$$

again under the assumption that the equilibrium solution is locally homogeneous, thus $\vec{\nabla} P_0 = 0 = \vec{\nabla} \times \vec{B}_0$. The pressure perturbation δP can further be related to the density perturbation $\delta\rho$ by means of the sound speed $c_{\rm s}$ of the neutral gas, $\delta P = c_{\rm s}^2 \delta\rho$.

As usual in a perturbation analysis, we decompose all of the perturbations, jointly represented by Q, into plane waves,

$$\delta Q \propto \mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x} - \omega t\right)} \ , \tag{4.292}$$

which turns the ideal magnetohydrodynamic equations into a set of algebraic equations. These are

$$\vec{k} \cdot \delta\vec{B} = 0 \ , \quad \omega\delta\vec{B} + \vec{k} \times \left(\delta\vec{v} \times \vec{B}_0\right) = 0 \tag{4.293}$$

for the magnetic field and

$$\omega\delta\rho - \rho_0 \vec{k} \cdot \delta\vec{v} = 0 \ , \quad \omega\delta\vec{v} - \frac{c_{\rm s}^2 \delta\rho}{\rho_0}\vec{k} + \frac{\left(\vec{k} \times \delta\vec{B}\right) \times \vec{B}_0}{4\pi\rho_0} = 0 \tag{4.294}$$

for the plasma velocity.

Without loss of generality, we can now rotate the coordinate frame such that $\vec{k}$ points along the positive x axis and that $\vec{B}_0$ falls into the x-y plane. Further, we denote the angle between $\vec{k}$ and $\vec{B}$ with ψ such that $\vec{k} \cdot \vec{B}_0 = kB_0 \cos\psi$. With this choice of coordinates, the vector products in (4.293) and (4.294) become

$$k \times \left(\delta\vec{v} \times \vec{B}_0\right) = kB_0 \begin{pmatrix} 0 \\ \delta v_y \cos\psi - \delta v_x \sin\psi \\ \delta v_z \cos\psi \end{pmatrix} \tag{4.295}$$

and

$$\left(\vec{k} \times \delta\vec{B}\right) \times \vec{B}_0 = kB_0 \begin{pmatrix} -\delta B_y \sin\psi \\ \delta B_y \cos\psi \\ \delta B_z \cos\psi \end{pmatrix} \tag{4.296}$$

Equations (4.293) for the magnetic field now specialise to

$$\delta B_x = 0 \ , \quad \delta B_y = \frac{kB_0}{\omega}\left(\delta v_x \sin\psi - \delta v_y \cos\psi\right) \ , \quad \delta B_z = -\frac{kB_0}{\omega}\delta v_z \cos\psi \ . \tag{4.297}$$

The continuity equation simplifies to

$$\frac{\delta\rho}{\rho_0} = \frac{k}{\omega}\delta v_x \tag{4.298}$$

and allows us to express the density fluctuation $\delta\rho$ by the velocity fluctuation δv_x. This, then, turns the three components of the Euler equations into

$$\delta v_x - \frac{k^2 c_\mathrm{s}^2}{\omega^2}\delta v_x - \frac{kB_0\delta B_y \sin\psi}{4\pi\rho_0\omega} = 0\,, \quad \delta v_{y,z} + \frac{kB_0\delta B_{y,z}\cos\psi}{4\pi\rho_0\omega} = 0\,. \quad (4.299)$$

?

Confirm the expressions given in (4.295) and (4.296) for the double vector products.

Next, we use the equations (4.297) for the magnetic field to eliminate the field fluctuations $\delta\vec{B}$ from the components (4.299) of the Euler equation. On the way, we introduce two velocities, the phase velocity $c_k = \omega/k$ of the plane-wave perturbations (4.292) and the so-called Alfvén velocity c_A through

$$c_\mathrm{A}^2 = \frac{B_0^2}{4\pi\rho_0}\,. \quad (4.300)$$

These definitions allow writing the three components of the Euler equation in the compact matrix form

$$\begin{pmatrix} c_k^2 - c_\mathrm{s}^2 - c_\mathrm{A}^2\sin^2\psi & c_\mathrm{A}^2\sin\psi\cos\psi & 0 \\ c_\mathrm{A}^2\sin\psi\cos\psi & c_k^2 - c_\mathrm{A}^2\cos^2\psi & 0 \\ 0 & 0 & c_k^2 - c_\mathrm{A}^2\cos^2\psi \end{pmatrix}\delta\vec{v} = 0\,. \quad (4.301)$$

Once the velocity perturbations are found from this propagation equation, the magnetic-field perturbations follow from (4.297), the density fluctuation from the continuity equation (4.298), and pressure fluctuations from the density fluctuations by multiplication with the squared sound speed. Since the density perturbations are caused exclusively by the x component of the velocity perturbations δv_x which point, by construction, into the direction of the wave vector, only longitudinal waves are responsible for the density fluctuations.

4.9.2 Alfvén waves

Let us focus on velocity perturbations in $\hat{e}_z$ direction first. For such perturbations, the propagation equation (4.301) requires the dispersion relation

$$c_k^2 = c_\mathrm{A}^2\cos^2\psi \quad (4.302)$$

or, since the phase velocity c_k of the plane-wave perturbation is $c_k = \omega/k$,

$$\omega = c_\mathrm{A} k\cos\psi = c_\mathrm{A}\vec{k}\cdot\hat{b}\,, \quad (4.303)$$

where the Alfvén speed occurs. Comparing with the ordinary sound speed in a gas, the expression (4.300) for the Alfvén speed is very intuitive: The squared Alfvén speed is the pressure of the magnetic field divided by the plasma density, just as the ordinary sound speed c_s is given by the ratio of the gas pressure and the gas density. The phase velocity of these Alfvén waves is

$$\frac{\omega}{k} = c_\mathrm{A}\cos\psi\,, \quad (4.304)$$

while their group velocity is

$$\frac{\partial\omega}{\partial\vec{k}} = c_\mathrm{A}\hat{b}\,. \quad (4.305)$$

We thus see that the Alfvén waves described by (4.303) are wave-like perturbations of the velocity field and the magnetic field transverse to their propagation direction *and* to the unperturbed magnetic field (Figure 4.10). Their group velocity has an absolute value depending only on the ratio of the magnetic pressure and the matter density. While the phase of the wave propagates along the wave vector $\vec{k}$ into the $\hat{e}_x$ direction, the group velocity points into the direction of the magnetic field. Alfvén waves thus transport physical quantities, for example their energy and momentum, along the magnetic field $\vec{B}$, independent of $\vec{k}$. The phase velocity of the Alfvén waves, $c_\mathrm{A}\cos\psi$, depends on the angle between $\vec{k}$ and $\vec{B}$ and vanishes if $\vec{k}$ is transverse to the magnetic field. Such Alfvén waves have a time-independent phase, and their energy propagates with the Alfvén velocity perpendicular to their wave vector. Alfvén wave packets, for example, with $\vec{k}\perp\vec{B}$ would propagate along $\vec{B}$, without changing their phase. If $\vec{k}$ and $\vec{B}$ are aligned, phase and group velocity become equal and point into the same direction.

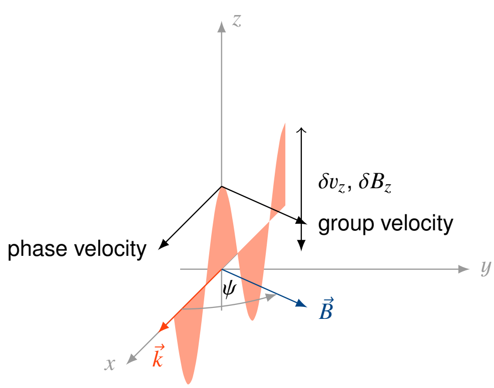

Figure 4.10 Alfvén waves are perturbations of the velocity field and the magnetic field perpendicular to their propagation direction $\vec{k}$ and the magnetic field $\vec{B}$.

Since the velocity perturbations δv_x and δv_y vanish for pure Alfvén waves, no density perturbations are associated with them, and the only component of the magnetic-field perturbation is

$$\delta B_z = -B_0\cos\psi\frac{\delta v_z}{c_k}\;. \tag{4.306}$$

The magnetic-field perturbation associated with Alfvén waves is thus antiparallel to the velocity perturbation, transverse to both the wave vector $\vec{k}$ and the magnetic field $\vec{B}_0$, and its amplitude is proportional to the component of the unperturbed magnetic field in the direction of the wave vector.

4.9.3 Slow and fast hydro-magnetic waves

Let us now consider waves described by the x and y components of the propagation equation (4.301),

$$\begin{pmatrix} c_k^2 - c_\mathrm{s}^2 - c_\mathrm{A}^2\sin^2\psi & c_\mathrm{A}^2\sin\psi\cos\psi \\ c_\mathrm{A}^2\sin\psi\cos\psi & c_k^2 - c_\mathrm{A}^2\cos^2\psi \end{pmatrix}\begin{pmatrix}\delta v_x\\ \delta v_y\end{pmatrix} = 0\;. \tag{4.307}$$

The dispersion relation is found requiring that the determinant of the coefficient matrix in this equation vanish, which gives a quadratic equation in the phase velocity c_k,

$$c_k^4 - c_k^2\left(c_A^2 + c_s^2\right) + c_A^2 c_s^2 \cos^2\psi = 0\,. \tag{4.308}$$

Its solutions are

$$c_{k,\pm}^2 = \frac{1}{2}\left[\left(c_A^2 + c_s^2\right) \pm \sqrt{\left(c_A^2 + c_s^2\right)^2 - 4c_A^2 c_s^2 \cos^2\psi}\,\right]\,. \tag{4.309}$$

Thus, a fast and a slow wave mode are possible (Figure 4.11). We analyse their modes in the special cases when the wave vector $\vec{k}$ is either aligned with the magnetic field $\vec{B}_0$, or transverse to it.

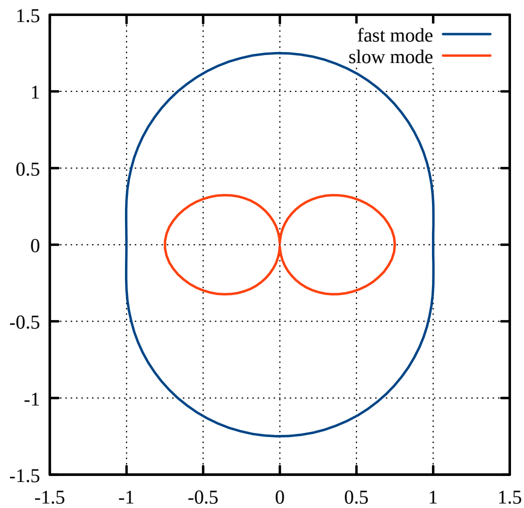

Figure 4.11 Polar velocity diagram of the fast and slow hydromagnetic modes. The polar angle is the angle ψ between the wave vector $\vec{k}$ and the magnetic field $\vec{B}$. The sound and Alfvén speeds are arbitrarily set to $c_s = 1$ and $c_A = 0.75$.

Let us begin with the case $\psi = 0$, when the perturbation propagates with or against the magnetic field. Then, $\cos\psi = 1$ and the dispersion relation becomes

$$c_k^2 = \frac{\omega^2}{k^2} = \frac{1}{2}\left(c_s^2 + c_A^2 \pm \left|c_s^2 - c_A^2\right|\right) = \begin{cases} c_s^2 & \text{or} \\ c_A^2 \end{cases}\,. \tag{4.310}$$

Accordingly, the fast wave propagates with the faster of the sound and the Alfvén velocities, the slow wave with the slower of these two. The propagation condition (4.301) shows that the wave travelling with the sound speed, the so-called acoustic mode, must have $\delta v_y = 0$ and is therefore logitudinal, while the wave travelling with the Alfvén speed, called the Alfvénic mode, is transversal since $\delta v_x = 0$. The acoustic mode creates density perturbations according to the continuity equation (4.298) while the Alfvénic mode does not because the density perturbations are proportional to δv_x. Similarly, the Alfvénic mode creates a transverse magnetic-field perturbation according to (4.297) while the acoustic mode has no magnetic-field perturbation associated.

For waves perpendicular to the magnetic field, $\vec{B} \perp \vec{k}$ and $\cos\psi = 0$, the phase velocities are

$$c_{k,\pm}^2 = \frac{1}{2}\left(c_s^2 + c_A^2 \pm c_s^2 + c_A^2\right) = \begin{cases} c_s^2 + c_A^2 & \text{or} \\ 0 \end{cases} \tag{4.311}$$

for the fast and the slow hydromagnetic waves. As for the Alfvén waves themselves, the phase velocity of the slow hydromagnetic waves then drops to zero. For $\psi = \pi/2$, the propagation condition (4.301) shows that the fast wave must be longitudinal while the slow wave must be transversal. Then, the fast wave creates density fluctuations and transversal magnetic-field perturbations as shown by (4.297), while the slow mode creates neither of them.

This concludes our brief introduction into the very rich field of magneto-hydrodynamics. Even neglecting any thermal motion of the plasma particles, viscosity or gravity, we found an interesting collection of phenomena, of which the Faraday rotation, the Alfvén waves, and the occurrence of the fast and the slow hydromagnetic waves were the most important.

Problems

1. Return to the ideal magneto-hydrodynamic equations (4.286), add a gravitational field, and assume a static, planar system infinitely extended in the x-y plane. Let the magnetic field be oriented parallel to the plane. For simplicity, assume further that the fluid is isothermal and that the ratio of the magnetic to the thermal pressure is constant.
   (a) Derive and solve an equation for the pressure as a function of distance z above the plane.
   (b) How is the magnetic field structured above the plane?

Suggested further reading: [2, 13, 15, 16, 18]

Chapter 5

Stellar Dynamics

5.1 The Jeans equations and Jeans' theorem

We begin this section by a derivation of the relaxation time scale (5.20), showing how long it takes a star orbiting through a system of stars to substantially change its velocity. In a way reminiscent of the derivation of the hydrodynamical equations, we then derive Jeans' equations (5.41) for the evolution of the number density and the mean velocity of stars in a stellar system. We then transform the Jeans equations to spherical polar coordinates and specialise them to stationary spherical systems in (5.56), showing that the main difference to hydrostatics is the anisotropy in velocity space. Next, we derive the tensor virial theorem (5.82) for stellar-dynamical systems and introduce Jeans' theorem.

5.1.1 Collision-less motion in a gravitational field

Particles in a gas or a fluid move almost unaccelerated until they meet another particle, which forces them to change their state of motion abruptly. As we have discussed before, hydrodynamics is based on the central assumption that the collisions occur on much smaller length scales λ than those macroscopic scales L that characterise the extent of the entire hydrodynamical system. In plasma physics, we had seen that the shielding of charges on the scale of the Debye length λ_D allows a hydrodynamical treatment despite the formally inifite range of electrostatic interactions, provided there are sufficiently many particles in the Debye volume $\approx \lambda_\mathrm{D}^3$. In all these cases, the interactions are effectively extremely short-ranged. Likewise, we had assumed in our treatment of local thermodynamical equilibrium in radiation transport that the mean free path of the photons be much smaller than the characteristic dimensions of the system under consideration.

Studying the motion of many point masses such as stars in a gravitational field, we encounter a fundamentally changed situation. The forces between the particles are now long-ranged and cannot be shielded. A single star in a galaxy, for instance, thus experiences not only the attraction of its nearest neighbours, but essentially the gravitational force exerted by all stars in the entire galaxy.

To give an illustrative example, let us consider a two-dimensional system, such as a galactic disk, which we shall assume to be infinitely extended for now and in whose centre we assume a star. The disk be randomly covered by stars in such a way that their mean number density is spatially constant (Figure 5.1).

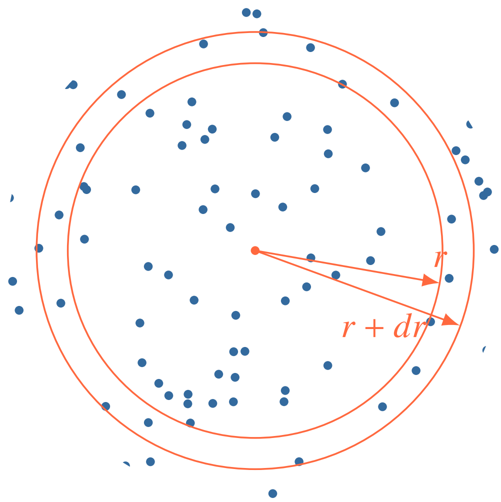

Figure 5.1 Illustration of a small section of a random star field, centered on an arbitrarily chosen star.

In a circular ring around the central star of radius r and width $\mathrm{d}r$, we find

$$\mathrm{d}N = 2\pi r \mathrm{d}r\, n \tag{5.1}$$

stars whose combined gravitational force on the central star is

$$\mathrm{d}F = 2\pi r \mathrm{d}r\, n \frac{Gm^2}{r^2} \tag{5.2}$$

if the mass m is assumed to be the same for all stars for simplicity. Of course, the directions of all forces cancel in the mean, but the contribution of arbitrarily distant rings diverges logarithmically,

$$\int \mathrm{d}F = 2\pi Gnm^2 \int \frac{\mathrm{d}r}{r} = 2\pi Gnm^2 \ln r\ . \tag{5.3}$$

Thus, the structure of the entire stellar system is important for the dynamics of the stars in the gravitational field.

In the spirit of the distinction between microscopic and macroscopic forces that we had made when introducing hydrodynamics, the forces in a system which is dominated by self-gravity are also macroscopic. Therefore, the collision terms, which describe the interaction on a microscopic scale, can be neglected here at least to first order of approximation. Thus, we begin our treatment of self-gravitating systems with the collision-less Boltzmann equation,

$$\mathrm{d}_t f(\vec{x}, \vec{v}, t) = \partial_t f + \dot{\vec{x}} \cdot \vec{\nabla} f + \dot{\vec{v}} \cdot \vec{\nabla}_{\vec{v}} f = 0 \tag{5.4}$$

5.1.2 The relaxation time scale

Before we turn to a detailed study of Eq. (5.4) in a gravitational field, we investigate approximately how the trajectory of a star through a galaxy which is composed of individual stars deviates from the trajectory through a hypothetical, "smooth" galaxy. We consider the passage of a star past another star employing Born's approximation, i.e. we integrate the deflection along a straight trajectory passing the deflecting star at an impact parameter b. The perpendicular force at the location x along the hypothetical, straight trajectory is

$$F_\perp = \left|-\vec{\nabla}_\perp \phi\right| = \left|-\frac{\partial}{\partial b}\frac{Gm^2}{\sqrt{b^2+x^2}}\right| = \frac{Gm^2 b}{(b^2+x^2)^{3/2}} , \tag{5.5}$$

where ϕ is the Newtonian gravitational potential. With $x \approx vt$, we have

$$F_\perp \approx \frac{Gm^2}{b^2}\left[1+\left(\frac{vt}{b}\right)^2\right]^{-3/2} , \tag{5.6}$$

and Newton's second law $m\dot{v}_\perp = F_\perp$ thus implies

$$\begin{aligned}\delta v_\perp &\approx \frac{Gm}{b^2}\int_{-\infty}^{\infty}\left[1+\left(\frac{vt}{b}\right)^2\right]^{-3/2}\mathrm{d}t \\ &= \frac{2Gm}{bv}\int_0^\infty \left(1+\tau^2\right)^{-3/2}\mathrm{d}\tau = \frac{2Gm}{bv} .\end{aligned} \tag{5.7}$$

?

How do you solve an integral like that in (5.7)?

Let N be the number of stars in the galaxy and R be its radius, then the fiducial test star experiences

$$\delta N = 2\pi b\delta b\, n = 2\pi b\delta b\frac{N}{\pi R^2} = \frac{2N}{R^2}b\delta b \tag{5.8}$$

such encounters with other stars at an impact parameter between b and $b+\delta b$. The mean quadratic velocity change is thus

$$\delta v_\perp^2 \approx \frac{2Nb\delta b}{R^2}\left(\frac{2Gm}{bv}\right)^2 = \frac{8NG^2m^2}{R^2v^2}\frac{\delta b}{b} . \tag{5.9}$$

Integrating this expression, we need to take into account that the assumption of Born's approximation requires that

$$\delta v_\perp \lesssim v \quad\Rightarrow\quad \frac{2Gm}{bv} \lesssim v \quad\Rightarrow\quad b \gtrsim b_{\mathrm{min}} = \frac{Gm}{v^2} , \tag{5.10}$$

and thus we obtain

$$\Delta v_\perp^2 = \int_{b_{\mathrm{min}}}^{\infty}\delta v_\perp^2 \approx 2N\left(\frac{2Gm}{Rv}\right)^2 \ln b\Big|_{b_{\mathrm{min}}}^{R} \equiv 2N\left(\frac{2Gm}{Rv}\right)^2 \ln\Lambda , \tag{5.11}$$

where

$$\ln\Lambda \equiv \ln\frac{R}{b_{\mathrm{min}}} = \ln\frac{Rv^2}{Gm} ; \tag{5.12}$$

is the so-called *Coulomb logarithm*. A typical velocity for the stars in a galaxy of mass $M = Nm$ is, according to the virial theorem,

$$v^2 \approx \frac{GMm}{R} \quad\Rightarrow\quad R \approx \frac{GNm}{v^2} . \tag{5.13}$$

Using this, we obtain

$$\frac{\Delta v_\perp^2}{v^2} \approx \frac{8 \ln \Lambda}{N} \ . \tag{5.14}$$

This shows by which relative amount the star's velocity is changed during one passage through the galaxy. The Coulomb logarithm $\ln \Lambda$ follows from

$$\ln \Lambda = \ln \frac{R}{b_{\min}} = \ln \frac{R v^2}{G m} \approx \ln N \ , \tag{5.15}$$

i.e. the relative velocity change is approximated by

$$\frac{\Delta v_\perp^2}{v^2} \approx \frac{8 \ln N}{N} \ . \tag{5.16}$$

After n_{cross} passages through the galaxy, the total relative velocity change will approximately be

$$n_{\text{cross}} \frac{8 \ln N}{N} \ . \tag{5.17}$$

For this expression to be of order unity, the number of passages needs to be

$$n_{\text{cross}} \approx \frac{N}{8 \ln N} \ . \tag{5.18}$$

Since one passage takes approximately the time

$$t_{\text{cross}} \approx \frac{R}{v} \ , \tag{5.19}$$

a substantial velocity change needs the *relaxation time*

$$t_{\text{relax}} \approx \frac{R}{v} \frac{N}{8 \ln N} \ . \tag{5.20}$$

Example: Relaxation of a galaxy

In a galaxy, we typically have a crossing time scale of

$$t_{\text{cross}} \approx \frac{10\,\text{kpc}}{200\,\text{km}\,\text{s}^{-1}} \approx 5 \cdot 10^7\,\text{yr} \tag{5.21}$$

and perhaps $N \approx 10^{11}$ stars. The relaxation time thus turns out to be

$$t_{\text{relax}} \approx 3 \cdot 10^{16}\,\text{yr} \ , \tag{5.22}$$

which is much more than the age of the Universe. This illustrates that in many, if not most astrophysically relevant systems, the collision-less Boltzmann equation can safely be used. ◂

5.1.3 The Jeans equations

The derivation of the Jeans equation, which will now follow, is formally similar to the derivation of the hydrodynamical equations. Yet, there are several important conceptual differences which justify going through the derivation again. First of all, we now have to do with a collection of individual, indentifiable

Example: Relaxation of a globular cluster

A counter-example is given by globular clusters. There, the number of stars is much smaller, $N \approx 10^5$, and crossing times are of order $t_{\rm cross} \approx 10^5$ yr. Their relaxation time scale is therefore

$$t_{\rm relax} \approx 10^8 \text{ yr} , \tag{5.23}$$

which is short compared to the life time of the globular cluster. In such cases, therefore, collisions do play an essential role. ◂

"particles", namely the stars in a stellar system, orbiting under their mutual gravitational interaction and possibly in an external, more or less smooth gravitational potential ϕ. Second, when we integrated over the momentum subspace during the derivation of hydrodynamics, we introduced an integral measure to ensure that the integral was relativistically invariant. There is no need to do so here since we can treat the stars in a stellar system as non-relativistically moving objects. Third, on the way to hydrodynamics, we introduced the four-vector J^μ for the particle-current density and the energy-momentum tensor $T^{\mu\nu}$ of the fluid and showed that the hydrodynamical equations followed from the vanishing four-divergences of J^μ and $T^{\mu\nu}$. Since we now have a collection of individual point masses, the introduction of continuous quantities such as the current densities of stars, momentum and energy is not necessarily justified. What we shall introduce, though, is the mean spatial number density $n(t, \vec{x})$ of the stars at the position $\vec{x}$ and at the time t.

We thus begin again with Boltzmann's collision-less equation, in which the right-hand side is set to zero, $\mathrm{d}_t f = 0$. Consider $f(\vec{x}, \vec{v}, t)$ as a function of position, velocity and time, and replace the time derivative of the velocity according to Newton's second law,

$$\dot{\vec{v}} = \frac{\vec{F}}{m} = -\vec{\nabla}\phi \tag{5.24}$$

to obtain

$$\partial_t f + \vec{v} \cdot \vec{\nabla} f - \vec{\nabla}\phi \cdot \vec{\nabla}_{\vec{v}} f = 0 . \tag{5.25}$$

Similar to the derivation of the hydrodynamical equations, we now form velocity moments of equation (5.25) by multiplying the collision-less Boltzmann equation with powers of the velocity and integrating over velocity space,

$$\partial_t \int \mathrm{d}^3 v\, f + \int \mathrm{d}^3 v\, \vec{v} \cdot \vec{\nabla} f - \vec{\nabla}\phi \cdot \int \mathrm{d}^3 v\, \vec{\nabla}_{\vec{v}} f = 0 . \tag{5.26}$$

The last term here simply gives boundary terms at infinity which vanish under the assumption that there are no infinitely fast point masses,

$$f(\vec{x}, \vec{v}, t) \to 0 \quad \text{for} \quad \left|\vec{v}\right| \to \infty . \tag{5.27}$$

In the second term, the gradient can be pulled out of the integral since it operates on the spatial coordinates $\vec{x}$, while the integration is carried out over the velocity. Equation (5.26) then gives

$$\partial_t n + \vec{\nabla} \cdot \int \mathrm{d}^3 v\, f \vec{v} = 0 . \tag{5.28}$$

Since the mean velocity is defined as

$$\langle \vec{v} \rangle = \frac{1}{n} \int \mathrm{d}^3 v \, f \vec{v} \,, \tag{5.29}$$

we find the continuity equation for our point masses,

$$\partial_t n + \vec{\nabla} \cdot (n \langle \vec{v} \rangle) = 0 \,, \tag{5.30}$$

as we might have expected. Notice in particular that we have introduced the mean spatial number density

$$n(\vec{x}, t) = \int \mathrm{d}^3 v \, f(\vec{x}, \vec{v}, t) \tag{5.31}$$

of the stars here. As an integral over the one-particle phase-space distribution f, this is a well-defined quantity, which should however not be confused with the smooth matter density of a fluid. Given the discrete nature of the stars in a stellar system, their spatial number density may fluctuate considerably. Moreover, it is not easily possible to move from the number density n to the matter density ρ by multiplying with a particle mass since the stars will typically have a wide mass distribution.

The second moment of Boltzmann's equation is taken by multiplying equation (5.25) with the velocity $\vec{v}$ prior to the integration over velocity space. In this way, further using that

$$\begin{aligned} \left(\vec{\nabla} f \cdot \vec{v}\right) \vec{v} &= \vec{\nabla} f^\top (\vec{v} \otimes \vec{v}) = \vec{\nabla} \cdot (f \vec{v} \otimes \vec{v}) \quad \text{and} \\ \left(\vec{\nabla}\phi \cdot \vec{\nabla}_{\vec{v}} f\right) \vec{v} &= \vec{\nabla}\phi^\top \left(\frac{\partial f}{\partial \vec{v}} \otimes \vec{v}\right) \end{aligned} \tag{5.32}$$

we obtain

$$\partial_t \int \mathrm{d}^3 v \, f \vec{v} + \vec{\nabla} \cdot \int \mathrm{d}^3 v \, f \vec{v} \otimes \vec{v} - \vec{\nabla}\phi^\top \int \mathrm{d}^3 v \left(\frac{\partial f}{\partial \vec{v}} \otimes \vec{v}\right) = 0 \,. \tag{5.33}$$

?

Verify the expressions (5.32) by your own calculation.

We continue by considering the third term, which can be integrated by parts to yield

$$\int \mathrm{d}^3 v \left(\frac{\partial f}{\partial \vec{v}} \otimes \vec{v}\right) = -\int \mathrm{d}^3 v \, f \, \frac{\partial \vec{v}}{\partial \vec{v}} = -n \mathbb{1}_3 \,, \tag{5.34}$$

if we can ignore boundary terms at infinity as before. This expression enables us to re-write (5.33) as

$$\partial_t (n \langle \vec{v} \rangle) + \vec{\nabla} \cdot (n \langle \vec{v} \otimes \vec{v} \rangle) + n \vec{\nabla}\phi = 0 \,, \tag{5.35}$$

where

$$\langle \vec{v} \otimes \vec{v} \rangle \equiv \frac{1}{n} \int \mathrm{d}^3 v \, f \vec{v} \otimes \vec{v} \tag{5.36}$$

is the velocity-dispersion tensor. This tensor can be re-written in terms of the velocity-correlation tensor and the average velocity components,

$$\langle \vec{v} \otimes \vec{v} \rangle = \langle (\vec{v} - \langle \vec{v} \rangle) \otimes (\vec{v} - \langle \vec{v} \rangle) \rangle + \langle \vec{v} \rangle \otimes \langle \vec{v} \rangle \equiv \sigma^2 + \langle \vec{v} \rangle \otimes \langle \vec{v} \rangle \,. \tag{5.37}$$

For convenience, we now substitute

$$\langle \vec{v} \rangle \to \vec{v} \tag{5.38}$$

since only averaged velocities and no velocities of individual particles remain. This allows us to write (5.35) as

$$\partial_t (n\vec{v}) + \vec{\nabla} \cdot \left(n\sigma^2\right) + \vec{\nabla} \cdot (n\vec{v} \otimes \vec{v}) + n\vec{\nabla}\phi = 0 \,. \tag{5.39}$$

Applying the product rule to the first and third terms and grouping terms conveniently, we can continue to write

$$\vec{v}\left[\partial_t n + \vec{\nabla} \cdot (n\vec{v})\right] + n\partial_t\vec{v} + \left(n\vec{v} \cdot \vec{\nabla}\right)\vec{v} + \vec{\nabla} \cdot \left(n\sigma^2\right) + n\vec{\nabla}\phi = 0 \,. \tag{5.40}$$

Noticing that the term in square brackets vanishes due to the continuity equation, we thus obtain the two equations

$$\partial_t n + \vec{\nabla} \cdot (n\vec{v}) = 0 \,,$$

$$\partial_t \vec{v} + \left(\vec{v} \cdot \vec{\nabla}\right)\vec{v} = -\vec{\nabla}\phi - \frac{1}{n}\vec{\nabla} \cdot \left(n\sigma^2\right) \,. \tag{5.41}$$

These are the *Jeans equations* which were derived for the first time by Maxwell, but first applied to stellar-dynamical problems by Sir James Jeans. As an equation of motion for the mean velocity components, the second equation corresponds to Euler's equation in ideal hydrodynamics, where the divergence of the tensor $n\sigma^2$ takes the role of the pressure gradient,

$$\frac{\vec{\nabla}P}{\rho} = \rho^{-1}\vec{\nabla} \cdot (P\mathbb{1}_3) \to \vec{\nabla} \cdot \left(n\sigma^2\right) \,. \tag{5.42}$$

5.1.4 Jeans equations in cylindrical and spherical coordinates

It is useful for many applications to write the distribution function f as a function not of Cartesian but of such coordinates that are adapted to the symmetry of a specific stellar-dynamical system under investigation. The Jeans equation then needs to be transformed from the Cartesian basis vectors $\{\hat{e}_x, \hat{e}_y, \hat{e}_z\}$ to those of the new, curvi-linear coordinates system. Let us carry out this transformation for the two frequent cases of cylindrical and spherical coordinates. Doing so, the transformations (3.209) and (3.214) derived earlier for the components of the acceleration need to be taken into account.

Since the gradient in cylindrical coordinates is

$$\vec{\nabla} = \hat{e}_r\partial_r + \frac{\hat{e}_\varphi}{r}\partial_\varphi + \hat{e}_z\partial_z \,, \tag{5.43}$$

we first obtain the components

$$a_r = -\partial_r\phi \,, \quad a_\varphi = -\frac{\partial_\varphi\phi}{r} \,, \quad a_z = -\partial_z\phi \tag{5.44}$$

of the equation of motion. With (3.209), we find

$$\dot{v}_r - \frac{v_\varphi^2}{r} = -\partial_r\phi \,, \quad \dot{v}_\varphi + \frac{v_r v_\varphi}{r} = -\frac{\partial_\varphi\phi}{r} \,. \tag{5.45}$$

This implies that the collision-less Boltzmann equation in cylindrical coordinates reads

$$\partial_t f + v_r \partial_r f + \frac{v_\varphi}{r}\partial_\varphi f + v_z \partial_z f + \left(\frac{v_\varphi^2}{r} - \partial_r \phi\right)\partial_{v_r} f - \left(\frac{v_r v_\varphi}{r} + \frac{\partial_\varphi \phi}{r}\right)\partial_{v_\varphi} f - \partial_z \phi \partial_{v_z} f = 0\ . \tag{5.46}$$

In spherical polar coordinates, we use the representation

$$\vec{\nabla} = \hat{e}_r \partial_r + \frac{\hat{e}_\theta}{r}\partial_\theta + \frac{\hat{e}_\varphi}{r\sin\theta}\partial_\varphi \tag{5.47}$$

of the gradient operator and transformation (3.214) of the acceleration components to find the fairly lengthy form

$$\partial_t f + v_r \partial_r f + \frac{v_\theta}{r}\partial_\theta f + \frac{v_\varphi}{r\sin\theta}\partial_\varphi f + \left[\frac{v_\theta^2 + v_\varphi^2}{r} - \partial_r \phi\right]\partial_{v_r} f - \left[\frac{v_r v_\theta}{r} - \frac{v_\varphi^2}{r}\cot\theta + \frac{\partial_\theta \phi}{r}\right]\partial_{v_\theta} f - \left[\frac{v_\varphi}{r}(v_r + v_\theta \cot\theta) + \frac{\partial_\varphi \phi}{r\sin\theta}\right]\partial_{v_\varphi} f = 0 \tag{5.48}$$

for the collision-less Boltzmann equation, whose physical meaning remains of course unchanged.

Whatever coordinates we choose, the zeroth moment of the collision-less Boltzmann equation must reproduce the continuity equation (5.30), with the appropriate representation of the divergence operator in the coordinate system chosen. Let us multiply the Boltzmann equation in spherical coordinates, (5.48), with the radial velocity component v_r and then integrate it over the complete velocity subspace of phase space. The result is the still lengthy expression

$$\partial_t \left(n\langle v_r\rangle\right) + \partial_r \left(n\left\langle v_r^2\right\rangle\right) + \frac{\partial_\theta}{r}\left(n\langle v_r v_\theta\rangle\right) + \frac{\partial_\varphi}{r\sin\theta}\left(n\left\langle v_r v_\varphi\right\rangle\right) - \frac{n}{r}\left\langle v_\theta^2 + v_\varphi^2\right\rangle + n\partial_r\phi + \frac{2n}{r}\left\langle v_r^2\right\rangle + \frac{n}{r}\langle v_r v_\theta\rangle\cot\theta = 0\ , \tag{5.49}$$

where we have kept the order and the arrangement of the terms like those from which they originate in (5.48). The decisive step in deriving (5.49) are partial integrations in velocity space.

?

Derive (5.49) yourself, following the steps described in the text.

5.1.5 Application to spherical systems

Equation (5.49) can be considerably simplified under the following natural assumptions. First, let us assume that the average velocities in the polar and azimuthal directions vanish,

$$\left\langle v_\varphi\right\rangle = 0 = \langle v_\theta\rangle\ . \tag{5.50}$$

Then, let us further assume that the velocity components are statistically independent of each other,

$$\langle v_r v_\theta\rangle = 0 = \left\langle v_r v_\varphi\right\rangle\ , \tag{5.51}$$

and that the situation is static, allowing us to ignore the partial time derivative. If we further introduce the velocity dispersions as the averages

$$\sigma^2_{r,\theta,\varphi} = \frac{1}{n} \int \mathrm{d}^3 v \, v^2_{r,\theta,\varphi} f \; , \tag{5.52}$$

we arrive at the much simpler equation

$$\partial_r \left(n\sigma_r^2 \right) + \frac{n}{r} \left[2\sigma_r^2 - \left(\sigma_\theta^2 + \sigma_\varphi^2 \right) \right] = -n\partial_r \phi \; . \tag{5.53}$$

Notice that we have neither used the continuity equation nor the explicit assumption of spherical symmetry here, but exclusively the first, radial moment of the collision-less Boltzmann equation together with an assumed isotropy in velocity space, expressed by the conditions (5.50) and (5.51).

Given this isotropy in velocity space, it is natural to assume that the polar and azimuthal velocity dispersions be equal,

$$\sigma_\theta^2 = \sigma_\varphi^2 \; . \tag{5.54}$$

We relate them to the radial velocity dispersion σ_r^2 by an anisotropy parameter β such that

$$\sigma_\theta^2 = \sigma_r^2 (1 - \beta) = \sigma_\varphi^2 \; . \tag{5.55}$$

Typically, the anisotropy parameter is non-negative, $\beta \geq 0$. If $\beta > 0$, radial motion dominates, while tangential motion dominates if $\beta < 0$. The anisotropy parameter itself cannot generally be assumed to be constant, but should be taken as depending on the radius r.

We are finally left with the radial Jeans equation

$$\partial_r \left(n\sigma_r^2 \right) + \frac{2\beta(r)}{r} n\sigma_r^2 = -n\partial_r \phi \; , \tag{5.56}$$

which is a first-order, linear, ordinary and inhomogeneous differential equation for the quantity $n\sigma_r^2$. It is easily solved by variation of constants. The general homogeneous solution is quickly found to be

$$n\sigma_r^2 = C \exp\left(-2 \int_0^r \frac{\beta(x)}{x} \mathrm{d}x \right) \; , \tag{5.57}$$

where the constant C is chosen such that $n\sigma_r^2 = C$ at the centre, $r = 0$, and x was introduced merely as a radial integration variable. For solving the inhomogeneous equation, we now allow C to vary with radius, $C = C(r)$. Then, (5.56) gives the differential equation

$$C'(r) = -n\partial_r \phi \exp\left(2 \int_0^r \frac{\beta(x)}{x} \mathrm{d}x \right) \tag{5.58}$$

for $C(r)$ since the exponential from (5.57) was constructed to solve the homogeneous equation (5.56) in the first place. This equation can formally be integrated to give

$$C(r) = \int_r^\infty \mathrm{d}y \left[n(y)(\partial_r \phi)(y) \exp\left(2 \int_0^y \frac{\beta(x)}{x} \mathrm{d}x \right) \right] \; , \tag{5.59}$$

where y was introduced as another radial integration variable and the boundary condition was chosen such that $C \to 0$ for $r \to \infty$ irrespective of what $\beta(r)$ may be. Returning with this result to (5.57), we obtain the solution

$$n\sigma_r^2 = \int_r^\infty \mathrm{d}y \left[n(y)(\partial_r\phi)(y) \exp\left(2\int_r^y \frac{\beta(x)}{x}\mathrm{d}x\right)\right] \tag{5.60}$$

for the radial velocity dispersion σ_r^2 times the stellar density n. For a spherically-symmetric system, we can further write the radial derivative of the gravitational potential as

$$\partial_r\phi = \frac{GM(r)}{r^2} , \tag{5.61}$$

which enables us to write

$$n\sigma_r^2 = G\int_r^\infty \mathrm{d}y \left[\frac{M(y)n(y)}{y^2} \exp\left(2\int_r^y \frac{\beta(x)}{x}\mathrm{d}x\right)\right] . \tag{5.62}$$

?

Carry out all steps yourself that lead from the radial Jeans equation (5.56) to the solution (5.62).

Many studies of stellar dynamics begin here. In principle, the radial stellar density $n(r)$ is observable through the surface brightness of an observed stellar system. By spectroscopy, the radial velocity dispersion σ_r^2 is accessible. If an anisotropy parameter $\beta(r)$ can now be reasonably guessed, (5.62) allows determining the mass (Figure 5.2).

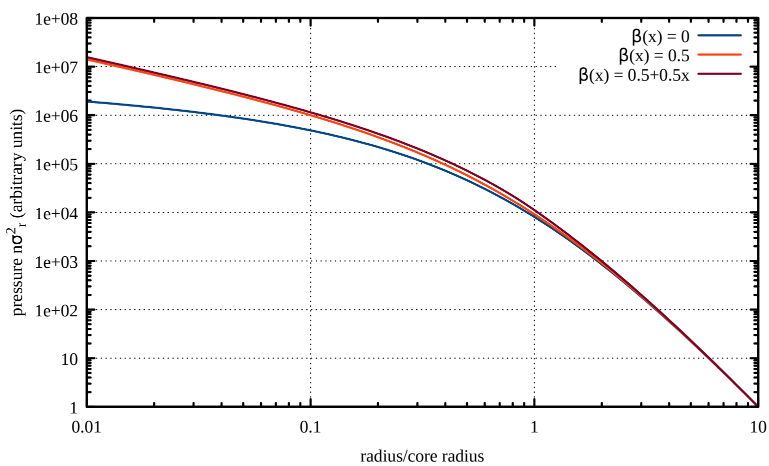

Figure 5.2 The effective kinematic pressure in a spherically-symmetric system according to Jeans' equation is shown as a function of the three-dimensional radius for different anisotropy parameters β.

The radial velocity dispersion multiplied with the stellar number density is of course not quite an observable quantity. Only the line-of-sight component of stellar velocities can typically be measured by the red- or blueshift of spectral lines. Since the red- and blueshifts of many stars generally appear superposed, lines appear broadened by the motion of the stars within the gravitational potential and shifted by the systemic velocity of the potential well as a whole relative to us as observers. The Doppler-broadened width of the spectral lines is the observable quantity to be measured, and it is directly related to the line-of-sight averaged velocity dispersion. Since the observed spectral line is a superposition of lines in the spectra of many stars, the observed line is dominated by those velocities that are represented by the most intense stellar light. Assuming that the stellar light is related to the stellar density by some constant factor, what we see is thus the density-weighted component of the stellar velocity along the line-of-sight.

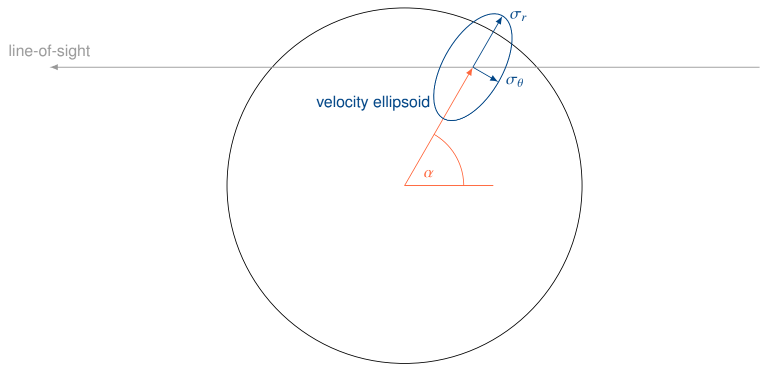

Figure 5.3 Illustration of the projection of the radial and tangential velocity dispersions along the line-of-sight.

Imagine a line-of-sight that passes through the spherical galaxy at a projected distance s from the line-of-sight through its centre at an azimuthal angle that we can without loss of generality assume to be zero, $\varphi = 0$. Further, let α be the angle between this line-of-sight and the radial direction. Then, the projected velocity component parallel to the line-of-sight is (Figure 5.3)

$$v_\parallel = v_r \cos\alpha + v_\theta \sin\alpha \,, \tag{5.63}$$

whose density-weighted, averaged square we see,

$$\sigma_\parallel^2 = \frac{\int_{-\infty}^{\infty} \mathrm{d}z\, n(s,z) \left\langle v_\parallel^2 \right\rangle}{\int_{-\infty}^{\infty} \mathrm{d}z\, n(s,z)} \,. \tag{5.64}$$

The denominator normalises the line-of-sight weighting with the stellar density. Since it is proportional to the surface brightness of the stellar light, we abbreviate it by $I(s)$. The integral along the z direction is conveniently converted into an integral in radial direction by noting that $r^2 = s^2 + z^2$ such that, at constant projected distance s, we have $r\mathrm{d}r = z\mathrm{d}z$. We can thus write the normalisation integral as

$$I(s) = \int_{-\infty}^{\infty} \mathrm{d}z\, n(s,z) = 2\int_s^{\infty} \frac{r\mathrm{d}r}{z}\, n(s,z) = 2\int_s^{\infty} \frac{r\mathrm{d}r\, n(r)}{\sqrt{r^2 - s^2}} \tag{5.65}$$

and the projected velocity dispersion as

$$\sigma_\parallel^2 = \frac{2}{I(s)} \int_s^{\infty} \frac{r\mathrm{d}r\, n(r)}{\sqrt{r^2 - s^2}} \left\langle (v_r \cos\alpha + v_\theta \sin\alpha)^2 \right\rangle \,. \tag{5.66}$$

The average is easily carried out. The mixed average $\langle v_r v_\theta \rangle$ vanishes due to our isotropy assumption (5.51) such that

$$\left\langle (v_r \cos\alpha + v_\theta \sin\alpha)^2 \right\rangle = \sigma_r^2 \cos^2\alpha + \sigma_\theta^2 \sin^2\alpha = \sigma_r^2 \left[\cos^2\alpha + (1-\beta)\sin^2\alpha\right] = \sigma_r^2 \left(1 - \beta \sin^2\alpha\right) \quad (5.67)$$

remains, where the anisotropy parameter β from (5.55) was inserted. By definition of the angle α, we can further substitute

$$\sin^2\alpha = \frac{s^2}{r^2} \quad (5.68)$$

in (5.67) and (5.66). This gives the relation

$$\sigma_\parallel^2 = \frac{2}{I(s)} \int_s^\infty \frac{r\mathrm{d}r\, n(r)\sigma_r^2}{\sqrt{r^2 - s^2}} \left(1 - \frac{\beta(r)s^2}{r^2}\right) \quad (5.69)$$

between the observable, density-weighted line-of-sight velocity dispersion $\sigma_\parallel^2$ and the radial velocity dispersion σ_r^2 (Figure 5.4).

?

Confirm the expression (5.69) for the line-of-sight velocity dispersion.

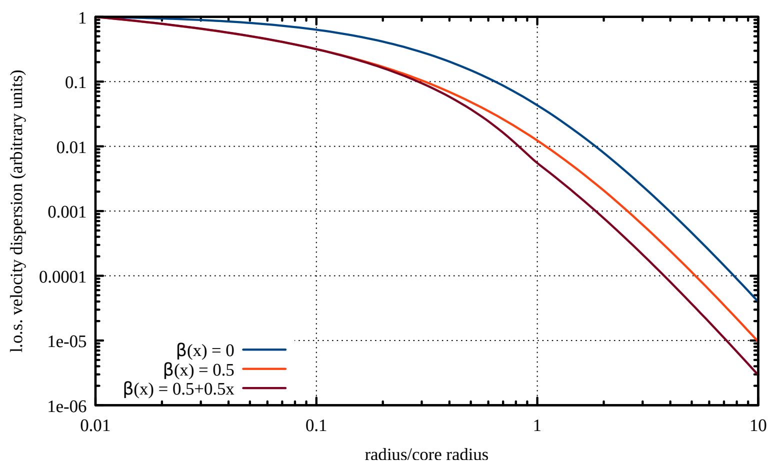

Figure 5.4 The observable line-of-sight velocity dispersion in a spherical system is shown as a function of projected radius for different anisotropy parameters β.

The anisotropy parameter can finally be eliminated by means of (5.56), since

$$\frac{\beta(r)}{r} n\sigma_r^2 = -\frac{1}{2}\left[\partial_r\left(n\sigma_r^2\right) + n\partial_r\phi\right] . \quad (5.70)$$

Inserting this expression into (5.69) and rearranging leads us to

$$I(s)\sigma_\parallel^2 - Gs^2 \int_s^\infty \frac{\mathrm{d}r\, n(r)M(r)}{r^2\sqrt{r^2 - s^2}} = \int_s^\infty \frac{r\mathrm{d}r}{\sqrt{r^2 - s^2}} \left[2n(r)\sigma_r^2 + \frac{s^2\partial_r(n\sigma_r^2)}{r}\right] . \quad (5.71)$$

This is an integro-differential equation for the density-weighted, radial velocity-dispersion profile $n\sigma_r^2$, determined by two observables, the surface-brightness profile $I(s)$ and the line-of-sight velocity dispersion $\sigma_\parallel^2$, together with a mass model $M(r)$. With (5.70), the solution can be used to constrain the anisotropy parameter $\beta(r)$.

5.1.6 The tensor virial theorem in stellar dynamics

We return to the Jeans equation in the form (5.35),

$$\partial_t \left(n \langle \vec{v} \rangle\right) + \vec{\nabla} \cdot \left(n \langle \vec{v} \otimes \vec{v} \rangle\right) + n\vec{\nabla}\phi = 0 \; . \tag{5.72}$$

Multiplication with the particle mass m and the spatial position vector $\vec{x}$, followed by integration over d^3x yields

$$\int d^3x\, \vec{x} \otimes \partial_t \left(\rho \langle \vec{v} \rangle\right) = -\int d^3x\, \vec{x} \otimes \vec{\nabla} \cdot \left(\rho \langle \vec{v} \otimes \vec{v} \rangle\right) - \int d^3x\, \vec{x} \otimes \rho\vec{\nabla}\phi \; . \tag{5.73}$$

We had seen already in (3.178) that the second term on the right-hand side is Chandrasekhar's tensor of the potential energy,

$$U = -\int d^3x\, \vec{x} \otimes \rho\vec{\nabla}\phi \; , \tag{5.74}$$

whose trace is the system's potential energy, as was shown in (3.180),

$$\mathrm{Tr}\, U = \frac{1}{2}\int d^3x\, \rho\phi \; . \tag{5.75}$$

Now we return to the first term on the right-hand side of the spatial integral (5.73), which we write as an integral over a complete divergence and a correction term,

$$\begin{aligned} \int d^3x\, \vec{x} \otimes \vec{\nabla} \cdot \left(\rho \langle \vec{v} \otimes \vec{v} \rangle\right) &= \int d^3x\, \vec{\nabla} \cdot \left(\rho \langle \vec{v} \otimes \vec{v} \rangle \otimes \vec{x}\right) \\ &\quad - \int d^3x\, \rho \langle \vec{v} \otimes \vec{v} \rangle \left(\vec{\nabla} \otimes \vec{x}\right) \; . \end{aligned} \tag{5.76}$$

By Gauss' law, the integral over the divergence is the surface integral over $\rho\langle \vec{v} \otimes \vec{v} \rangle \otimes \vec{x}$, which vanishes if the surface completely encloses the system such that the density vanishes there. The remaining term is related to the tensor K of the kinetic energy,

$$\int d^3x\, \rho \langle \vec{v} \otimes \vec{v} \rangle \left(\vec{\nabla} \otimes \vec{x}\right) = \int d^3x\, \rho \langle \vec{v} \otimes \vec{v} \rangle = 2K \; , \tag{5.77}$$

whose trace is the total kinetic energy of the system. By means of the velocity-correlation tensor σ^2 defined in (5.37), we can split up the tensor of kinetic energy into a part T due to the bulk motion of the system, and another part Π due to the random motion of the stars about the mean motion. Specifically, we define

$$K = \frac{1}{2}T + \frac{1}{2}\Pi \; , \tag{5.78}$$

where T and Π are defined by

$$T \equiv \int d^3x\, \rho \langle \vec{v} \rangle \langle \vec{v} \rangle \; , \quad \Pi \equiv \int d^3x\, \rho\sigma^2 \; . \tag{5.79}$$

Quite evidently, the tensor T corresponds to the stress-energy tensor in ideal hydrodynamics up to the pressure term, while the tensor Π describes the momentum transport by unordered motion and is thus represents the pressure.

On the left-hand side of the spatial integral (5.73), the term

$$\int \mathrm{d}^3x\,\vec{x}\otimes\partial_t\left(\rho\left\langle\vec{v}\right\rangle\right) \tag{5.80}$$

remains. We symmetrise it by bringing it into the form

$$\frac{1}{2}\int \mathrm{d}^3x\,\left[\vec{x}\otimes\partial_t\left(\rho\left\langle\vec{v}\right\rangle\right)+\partial_t\left(\rho\left\langle\vec{v}\right\rangle\right)\otimes\vec{x}\right] \tag{5.81}$$

which, by comparison with (3.184), equals the second absolute time derivative of the inertial tensor I. We thus obtain the tensor virial theorem for collision-less systems,

$$\frac{1}{2}\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = T + \Pi + U\ . \tag{5.82}$$

?

Why is it appropriate to symmetrise the tensor given by (5.80)?

Taking the trace of this equation leads us back to the ordinary (scalar) virial theorem, if the mass distribution is static,

$$\frac{\mathrm{d}^2\,\mathrm{Tr}\,I}{\mathrm{d}t^2} = 0 \quad\Rightarrow\quad \mathrm{Tr}\,T + \mathrm{Tr}\,\Pi + \mathrm{Tr}\,U = 0\ . \tag{5.83}$$

Now, the sum of the traces of T and Π is twice the trace of the total kinetic-energy tensor,

$$\mathrm{Tr}\,T + \mathrm{Tr}\,\Pi = 2\,\mathrm{Tr}\,K = \int \mathrm{d}^3x\,\rho v^2\ , \tag{5.84}$$

and thus twice the total kinetic energy K, while $\mathrm{Tr}\,U$ is the total potential energy, as we have seen before. Thus,

$$2\,\mathrm{Tr}\,K = -\,\mathrm{Tr}\,U\ , \tag{5.85}$$

which is the ordinary scalar virial theorem.

5.1.7 Jeans' theorem

An integral of the motion in any field of force is any function $Q(\vec{x},\vec{v})$ of the phase-space coordinates that satisfies

$$\frac{\mathrm{d}Q\left(\vec{x},\vec{v}\right)}{\mathrm{d}t} = 0 \tag{5.86}$$

along all possible particle trajectories $[\vec{x}(t),\vec{v}(t)]$ through phase space. Integrals of the motion should not be confused with constants of the motion, which are less strongly defined as quantities that do not depend on time along one particular orbit. Any integral of the motion turns into a constant of the motion when evaluated along a particular orbit, but the reverse is not generally true.

An orbit of a classical particle in a Hamiltonian system always has six constants of the motion. Namely, let the orbit be specified by $\vec{x}(t)$ and $\vec{v}(t)$, then it can be uniquely traced back to an initial phase-space point $(\vec{x}_0, \vec{v}_0)$ by means of the equations of motion. These six numbers are constants of the motion, since they are independent of time along any trajectory.

For Hamiltonian systems, a potential $\phi(\vec{x})$ exists and the Hamiltonian equations of motion require $\dot{\vec{v}} = -\vec{\nabla}\phi(\vec{x})$. The condition (5.86) for Q to be an integral of the motion can then be cast into the form

$$\frac{\mathrm{d}Q\,(\vec{x},\vec{v})}{\mathrm{d}t} = \dot{\vec{x}} \cdot \vec{\nabla}Q + \dot{\vec{v}} \cdot \frac{\partial Q}{\partial \vec{v}} = \vec{v} \cdot \vec{\nabla}Q - \vec{\nabla}\phi \cdot \frac{\partial Q}{\partial \vec{v}} = 0\,. \tag{5.87}$$

By comparison with the collision-less Boltzmann equation, we see that Q is an integral of the motion if and only if it is a stationary solution of the collision-less Boltzmann's equation, i.e. a solution satisfying

$$\frac{\partial Q}{\partial t} = 0\,. \tag{5.88}$$

This leads us to *Jeans' theorem*:

Any stationary solution of the collision-less Boltzmann equation depends on the phase-space coordinates only through integrals of the motion, and conversely any function depending only on integrals of the motion is a stationary solution of the collision-less Boltzmann equation.

The *proof* of the first statement has already been given: If Q is a stationary solution of the collision-less Boltzmann equation it is by itself an integral of the motion. Regarding the second statement, let I_i, $1 \le i \le n$ be an arbitrary number n of integrals of the motion, and let $Q(I_1, I_2, \ldots, I_n)$ an arbitrary function exclusively depending on these integrals. Then,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial Q}{\partial I_i}\frac{\mathrm{d}I_i}{\mathrm{d}t} = 0\,, \tag{5.89}$$

and Q solves the collision-less Boltzmann equation.

Jeans' theorem is important because it guides the construction of physically meaningful phase-space densities. For example, in a static, spherically symmetric potential with isotropic orbits, the phase-space density can only be a function of the energy $E = \vec{v}^2/2 + \phi(r)$. If the potential remains spherically symmetric, but the orbits become anisotropic, the phase-space density will depend on E and the absolute value of the angular momentum L. In a static system with axial symmetry, the energy E and the component L_z of the angular momentum along the symmetry axis will be integrals of the motion, and the phase-space density will depend only on those.

One distinguishes isolating and non-isolating integrals of the motion. An isolating integral of the motion defines a subspace of phase space with dimension lowered by one to which an orbit is confined. Let I_1 be a first isolating integral of the motion, for example the energy. In the originally six-dimensional phase space, I_1 defines a five-dimensional subspace from which no orbit can escape. An further isolating integral I_2 will confine orbits to a four-dimensional subspace, and so forth. Isolating integrals such as the energy E or the angular-momentum vector $\vec{L}$ constrain the orbits. If n isolating integrals exist, orbits must be confined to a $(6-n)$-dimensional subspace of phase space. Isolating integrals are extraordinarily important while non-isolating integrals have no practical importance for stellar dynamics.

Orbits are called regular if they have as many isolating integrals as there are spatial dimensions; otherwise they are called irregular. Regular orbits in d

Example: Harmonic oscillator

The example of a harmonic oscillator in one spatial dimension may perhaps be instructive. Its phase space is two-dimensional, its energy is conserved. The constant energy confines the phase-space orbits of the oscillator to one-dimensional subspaces of phase space which are the ellipses defined by

$$E = \frac{m}{2}\left(\vec{v}^2 + \omega^2 x^2\right) = \text{const} . \tag{5.90}$$

Any one-dimensional harmonic oscillator must remain on the ellipse defined by its energy, and no harmonic oscillator with another energy will ever enter that subspace of phase space. This illustrates the isolating effect of the energy in phase space. ◀

spatial dimensions are thus confined to $2d - d = d$-dimensional subspaces of phase space. The one-dimensional harmonic oscillator is one example for a system with regular orbits.

Problems

1. A convenient model density profile for different kinds of astrophysical objects is the Hernquist profile

$$\rho(x) = \frac{\rho_0}{x(1+x)^3} , \tag{5.91}$$

   proposed by L. Hernquist (1990). The dimension-less radius x is defined as $x = r/a$, with a scale or core radius a.

   (a) Write the density amplitude ρ_0 in terms of the total mass M contained in the profile (5.91).
   (b) Derive the Newtonian potential $\phi(x)$ of objects with the Hernquist density profile.
   (c) Assuming that the number-density $n(x)$ of the stars in a Hernquist-like object follows the matter-density profile, and assuming that the anisotropy parameter $\beta = 1/2$ is independent of the radius, solve (5.62) for the radial velocity dispersion.
   (d) Calculate the profile (5.69) of the observable, line-of-sight velocity dispersion profile.

2. For the Hernquist profile (5.91), calculate Chandrasekhar's tensor $U^i{}_j$ of the potential energy.

5.2 Equilibrium and Stability

This section discusses issues of equilibrium and stability of self-gravitating systems. The isothermal sphere (5.111) is introduced first as a simple example for a solution of the static Jeans equation in spherical symmetry.

Equilibrium considerations are briefly mentioned, emphasising that self-gravitating systems have no stable equilibrium state. A linear perturbation analysis reveals the close analogy (5.130) between perturbations of the gravitational potential in a stellar-dynamical system and the longitudinal dielectricity in a plasma. The Jeans wave number (5.136) is derived as the boundary between stable and unstable perturbations. A detour on two-dimensional, self-gravitating systems leads to the solution (5.159) of Poisson's equation for a disk with a given surface-mass density. Finally, the dispersion relation (5.185) for linear perturbations of disks is derived, leading to Toomre's stability criterion (5.193) for disks.

5.2.1 The Isothermal Sphere

By Noether's theorem, spherical systems which are independent of time have orbits with at least the four integrals of the motion, which are the energy E and the angular momentum $\vec{L}$. Jeans' theorem then tells us that any (non-negative) function $f(E, \vec{L})$ of these integrals of the motion is a stationary solution of the collision-less Boltzmann equation and may thus represent a stable, self-gravitating system. Generally, the gravitational potential generated by a system with a phase-space distribution function f is determined by Poisson's equation,

$$\vec{\nabla}^2\phi = 4\pi G\rho = 4\pi Gm \int \mathrm{d}^3v\, f \,, \tag{5.92}$$

where m is the particle mass, assumed to be the same for all particles. If the system is also isotropic in velocity space, the phase-space density cannot depend on the direction of the angular momentum either. Then, the phase-space density may be taken to be a function of E and the absolute value $L = m|\vec{x} \times \vec{v}|$ of the angular momentum only,

$$f\left(E, \vec{L}\right) = f(E, L) \,. \tag{5.93}$$

Writing the Laplacian operator in spherical symmetry,

$$\vec{\nabla}^2 = \frac{1}{r^2}\partial_r\left(r^2\partial_r\right) \,, \tag{5.94}$$

the equation for the gravitational potential

$$\frac{1}{r^2}\partial_r\left(r^2\partial_r\phi\right) = 4\pi Gm \int \mathrm{d}^3v\, f\left(\frac{mv^2}{2} + m\phi, m\left|\vec{x} \times \vec{v}\right|\right) \tag{5.95}$$

follows as the fundamental equation for self-gravitating spherical systems in equilibrium.

It is now convenient to re-scale the gravitational potential ϕ and the energy $E = mv^2/2 + m\phi$ by subtracting a constant potential ϕ_0 and defining the shifted potential ψ and the shifted specific energy $\mathcal{E}$,

$$\psi \equiv -\phi + \phi_0 \,, \quad \mathcal{E} \equiv -\frac{E}{m} + \phi_0 = \psi - \frac{v^2}{2} \,. \tag{5.96}$$

Let us consider a simple example specified by a phase-space distribution function f entirely independent of L and depending exponentially on the shifted specific energy $\mathcal{E}$,

$$f(\mathcal{E}) = \frac{\tilde{n}}{(2\pi\sigma^2)^{3/2}} \mathrm{e}^{\mathcal{E}/\sigma^2} = \frac{\tilde{n}}{(2\pi\sigma^2)^{3/2}} \exp\left(\frac{\psi - v^2/2}{\sigma^2}\right) , \tag{5.97}$$

where the constant $\tilde{n}$ appears for normalisation. Integration over all velocities yields the number density n of the particles,

$$\int \mathrm{d}^3 v\, f(\mathcal{E}) = \frac{4\pi\tilde{n}\mathrm{e}^{\psi/\sigma^2}}{(2\pi\sigma^2)^{3/2}} \int_0^\infty \mathrm{d}v\, v^2 \mathrm{e}^{-v^2/(2\sigma^2)} = \tilde{n}\mathrm{e}^{\psi/\sigma^2} = n , \tag{5.98}$$

as must be the case for all phase-space densities. Poisson's equation for this system then reads

$$\frac{1}{r^2}\partial_r \left(r^2 \partial_r \psi\right) = -4\pi G n m = -4\pi G m \tilde{n} \mathrm{e}^{\psi/\sigma^2} . \tag{5.99}$$

Eliminating the re-scaled potential ψ and the exponential by means of (5.98),

$$\psi = \sigma^2 \ln \frac{n}{\tilde{n}} = \sigma^2 (\ln n - \ln \tilde{n}) , \quad \mathrm{e}^{\psi/\sigma^2} = \frac{n}{\tilde{n}} , \tag{5.100}$$

and substituting

$$\partial_r \psi = \sigma^2 \partial_r \ln n , \tag{5.101}$$

we can turn Poisson's equation (5.99) into an equation for the spatial number density n,

$$\frac{1}{r^2}\partial_r \left(r^2 \partial_r \ln n\right) = -\frac{4\pi G m}{\sigma^2} n , \tag{5.102}$$

which can of course also be considered as an equation for the mass density $\rho = nm$.

In ideal hydrodynamics, we had derived the equation (3.267),

$$M(r) = -\frac{r k_\mathrm{B} T}{mG} \left(\frac{\mathrm{d} \ln \rho_\mathrm{gas}}{\mathrm{d} \ln r} + \frac{\mathrm{d} \ln T}{\mathrm{d} \ln r}\right) \tag{5.103}$$

for a spherical gas mass in hydrostatic equilibrium with the gravitational-potential well given by its mass $M(r)$ enclosed by the radius r. If this gas is isothermal, $\mathrm{d}T/\mathrm{d}r = 0$, we can re-write (5.103) as

$$r^2 \partial_r \ln \rho_\mathrm{gas} = -\frac{mG}{k_\mathrm{B} T} \int_0^r \mathrm{d}r'\, r'^2 \rho_\mathrm{gas}(r') \tag{5.104}$$

if we consider the mass as being only contributed by the gas without any dark matter. Differentiating (5.104) with respect to r and dividing by r^2 yields

$$\frac{1}{r^2}\partial_r \left(r^2 \partial_r \ln \rho_\mathrm{gas}\right) = -\frac{4\pi G m}{k_\mathrm{B} T} \rho_\mathrm{gas} . \tag{5.105}$$

This equation is identical with our result (5.102) which we had previously derived from Jeans' theorem if we identify the velocity dispersion σ with the specific thermal energy,

$$\sigma^2 = \frac{k_\mathrm{B} T}{m} . \tag{5.106}$$

Thus, the corresponding self-gravitating, stellar-dynamical model with constant velocity dispersion σ^2 is called the *isothermal sphere*. The mean-squared velocity in the isothermal sphere is

$$\left\langle v^2 \right\rangle = \frac{1}{n} \int \mathrm{d}^3 v \, v^2 f = \frac{\int \mathrm{d}v \, v^4 \exp\left(\frac{-v^2}{2\sigma^2}\right)}{\int \mathrm{d}v \, v^2 \exp\left(\frac{-v^2}{2\sigma^2}\right)} = 3\sigma^2 \; . \tag{5.107}$$

Since no direction is preferred due to the spherical symmetry, the three individual velocity components thus have the same mean square

$$\left\langle v_x^2 \right\rangle = \left\langle v_y^2 \right\rangle = \left\langle v_z^2 \right\rangle = \sigma^2 \; . \tag{5.108}$$

One solution of the equation (5.102) for the density of the isothermal sphere can be obtained with the *ansatz*

$$n = C r^{-\alpha} \; . \tag{5.109}$$

The operations on the left-hand side of (5.102) yield

$$-\frac{\alpha}{r^2} \partial_r \left(r^2 \partial_r \ln r \right) = -\frac{\alpha}{r^2} \; . \tag{5.110}$$

Since the right-hand side scales with the radius as r^{-2}, the two sides can equal if and only if $\alpha = 2$. Therefore, the *ansatz* (5.109) is indeed a solution of (5.102) if $\alpha = 2$ and $C = \sigma^2/(2\pi G m)$, giving the matter density

$$\rho(r) = m n(r) = \frac{\sigma^2}{2\pi G r^2} \; . \tag{5.111}$$

This solution is called the *singular isothermal sphere*. It has the considerable advantage that the velocity of test particles on circular orbits around its centre is independent of radius,

$$v_{\mathrm{circ}}^2 = \frac{G M(r)}{r} = 4\pi G \int_0^r r^2 \mathrm{d}r \, \rho(r) = 2\sigma^2 \; , \tag{5.112}$$

which is observed in the vast majority of galaxies. Besides the central singularity, a substantial conceptual disadvantage is that its mass grows linearly with the radius and is thus formally infinite. Of course, this is an inevitable consequence of the assumption that the gas is isothermal: If this is so, the gas distribution must extend to infinity.

Another solution of (5.102) which avoids the central singularity can be found numerically. For doing so, we conveniently introduce the dimension-less quantities

$$x \equiv \frac{r}{r_0} \; , \quad y \equiv \frac{\rho}{\rho_0} \; , \tag{5.113}$$

where ρ_0 is meant to be the finite central density. Then, the equation for the scaled density y is

$$\partial_x \left(x^2 \partial_x \ln y \right) = -\frac{4\pi G}{\sigma^2} \rho_0 r_0^2 \, y x^2 \; . \tag{5.114}$$

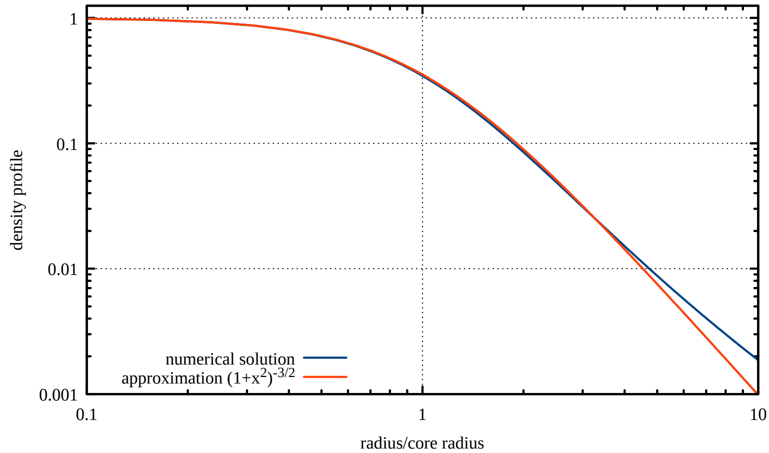

Figure 5.5 The density profile of the non-singular isothermal solution is shown together with its approximation.

If we define the scale radius r_0 to be related to the central density and the velocity dispersion by

$$r_0 \equiv \sqrt{\frac{9\sigma^2}{4\pi G\rho_0}} \,, \tag{5.115}$$

equation (5.114) simplifies to

$$\partial_x \left(x^2 \partial_x \ln y\right) = -9yx^2 \,, \tag{5.116}$$

which can numerically be integrated with the appropriate boundary conditions

$$y(0) = 1 \,, \quad \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_0 = 0 \,. \tag{5.117}$$

?

Compare (5.116) with the Lane-Emden equation (3.259) and the scale radius r_0 from (5.115) with r_0 from (3.258). Should they be related, and if so, for which n?

These boundary conditions mean that the central density is indeed ρ_0 and that the density profile is flat at the centre. For sufficiently small radii, the numerical result is very well approximated by (Figure 5.5)

$$y(x) \approx \left(1 + x^2\right)^{-3/2} \,. \tag{5.118}$$

For $x \lesssim 4.5$ or $r \lesssim 4.5r_0$, the relative deviation between the true numerical solution of (5.116) and the approximate solution (5.118) is $\lesssim 10\%$. As expected from isothermality, the total mass of the non-singular isothermal sphere still diverges. Since the density falls off asymptotically like r^{-3} for $r \gg r_0$, the mass must diverge logarithmically for $r \to \infty$.

5.2.2 Equilibrium and Relaxation

Is there an equilibrium state of a self-gravitating system, which corresponds to an entropy maximum? The entropy

$$S \propto -\int_{\text{phase space}} \mathrm{d}^3x\mathrm{d}^3p\, p \ln p \tag{5.119}$$

is maximised if and only if p is the distribution function of the isothermal sphere. However, the isothermal sphere has infinite mass and energy and can thus not be an exact description of a thermodynamical equilibrium state. This implies that there is no thermodynamical equilibrium of a self-gravitating system, and that self-gravitating systems cannot have stable final configurations, but at best long-lived transient states!

If we populate a narrow region in phase space with N stars, their orbits will have slightly different initial conditions. As time proceeds, they will progressively evolve away from each other and thus occupy a growing part of phase space. This *phase mixing* causes the averaged phase-space distribution $\bar{f}$ to decrease, because the averaged phase-space density is progressively diluted. Thus, the *macroscopic entropy*

$$\bar{S} \propto -\int \mathrm{d}^3x\mathrm{d}^3v\,\bar{f}\ln\bar{f} \tag{5.120}$$

does indeed increase.

This process of phase mixing is in fact hardly different from the thermodynamical trend to equilibrium. There, too, the increase of entropy is caused by macroscopically averaging over processes which are otherwise reversible. If the potential is changed while the particles are moving through it, energy can be transported from particles to others. If, for example, the system contracts while a star approaches its centre, the potential deepens and the star looses energy. Other stars can gain considerable amounts of energy; this process is called *violent relaxation* (Lynden-Bell).

5.2.3 Linear analysis and the Jeans swindle

In a way very similar to the derivation of the dielectricity tensor in plasma physics, we now consider an equilibrium solution f_0, ϕ_0 of the coupled system of the collision-less Boltzmann equation and the Poisson equation,

$$\begin{aligned} \frac{\partial f}{\partial t} + \vec{v}\cdot\vec{\nabla}f - \vec{\nabla}\phi\cdot\frac{\partial f}{\partial\vec{v}} &= 0\,, \\ \vec{\nabla}^2\phi &= 4\pi Gm\int\mathrm{d}^3v\,f\,. \end{aligned} \tag{5.121}$$

In a stationary equilibrium state, $\partial f_0/\partial t = 0$. As usual for a linear stability analysis, we perturb f_0 and ϕ_0 by small amounts δf and $\delta\phi$ and linearise the equations in these perturbations. The result is

$$\begin{aligned} \frac{\partial\delta f}{\partial t} + \vec{v}\cdot\vec{\nabla}\delta f - \vec{\nabla}\phi_0\cdot\frac{\partial\delta f}{\partial\vec{v}} - \vec{\nabla}\delta\phi\cdot\frac{\partial f_0}{\partial\vec{v}} &= 0\,, \\ \vec{\nabla}^2\delta\phi &= 4\pi Gm\int\mathrm{d}^3v\,\delta f\,. \end{aligned} \tag{5.122}$$

Poisson's equation implies one peculiarity here which needs to be emphasised. Suppose we adopt as the unperturbed equilibrium state an infinitely extended, homogeneous phase-space distribution f_0, which implies a constant density ρ_0 and a potential ϕ_0 given by

$$\vec{\nabla}^2\phi_0 = 4\pi G\rho_0\,. \tag{5.123}$$

Due to the infinite extent of this matter distribution and its homogeneity, we must have

$$\vec{\nabla}\phi_0 = 0 \tag{5.124}$$

because there cannot be any gravitational force on a test particle in a surrounding homogeneous matter density. This condition complies with the Poisson equation if and only if $\rho_0 = 0$. An infinitely extended, homogeneous matter distribution is possible in Newtonian gravity only if it has no matter; it is therefore generally inconsistent with Newtonian gravity. The reason is profound: In Newtonian gravity, the boundary conditions for the potential are of decisive importance for the solution of the Poisson equation, but an infinitely extended mass distribution has no boundary. Only in General Relativity, this problem is satisfactorily solved.

We rather invoke the "Jeans swindle" and set $\phi_0 = 0$. This is practically permissible if we study perturbations whose spatial scales are small compared to possible scales in the smooth background density ρ_0. By the "Jeans swindle", we simply ignore the potential ϕ_0 and obtain the linearised equations

$$\begin{aligned} \frac{\partial \delta f}{\partial t} + \vec{v}\cdot\vec{\nabla}\delta f - \vec{\nabla}\delta\phi\cdot\frac{\partial f_0}{\partial \vec{v}} &= 0\ , \\ \vec{\nabla}^2\delta\phi &= 4\pi G m \int \mathrm{d}^3 v\, \delta f\ . \end{aligned} \tag{5.125}$$

Now, we decompose the spatial and temporal dependence of the perturbations into plane waves,

$$\delta f = \delta\hat{f}\, \mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)}\ , \quad \delta\phi = \delta\hat{\phi}\, \mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)}\ , \tag{5.126}$$

where the amplitude $\delta\hat{f}$ of the phase-space distribution function will generally still depend on the velocity coordinates of phase space. The perturbation amplitudes must then satisfy the equations

$$\begin{aligned} -\mathrm{i}\omega\delta\hat{f} + \mathrm{i}\vec{v}\cdot\vec{k}\delta\hat{f} - \mathrm{i}\delta\hat{\phi}\vec{k}\cdot\frac{\partial f_0}{\partial\vec{v}} &= 0\ , \\ -k^2\delta\hat{\phi} &= 4\pi G m \int \mathrm{d}^3 v\, \delta\hat{f}\ . \end{aligned} \tag{5.127}$$

The perturbed Boltzmann equation can be solved to relate the perturbation of the phase-space distribution $\delta\hat{f}$ to the potential perturbation $\delta\hat{\phi}$,

$$\delta\hat{f} = \delta\hat{\phi}\,\vec{k}\cdot\frac{\partial f_0}{\partial\vec{v}}\,\frac{1}{\vec{k}\cdot\vec{v}-\omega}\ , \tag{5.128}$$

which can in turn be inserted into the perturbed Poisson equation to find

$$-k^2\delta\hat{\phi} = 4\pi G m \int \mathrm{d}^3 v\,\vec{k}\cdot\frac{\partial f_0}{\partial\vec{v}}\,\frac{\delta\hat{\phi}}{\vec{k}\cdot\vec{v}-\omega}\ . \tag{5.129}$$

Since the potential perturbation $\delta\hat{\phi}$ does not depend on $\vec{v}$, (5.129) shows that non-vanishing perturbations $\delta\hat{\phi} \neq 0$ are possible only if

$$\boxed{1 + \frac{4\pi G m}{k^2}\int \mathrm{d}^3 v\,\vec{k}\cdot\frac{\partial f_0}{\partial\vec{v}}\,\frac{1}{\vec{k}\cdot\vec{v}-\omega} = 0}\ . \tag{5.130}$$

This correponds exactly to the longitudinal dielectricity $\hat{\varepsilon}_{\|}$ (4.69) from plasma physics,

$$\hat{\varepsilon}_{\|} = 1 - \frac{4\pi e^2}{k^2} \int \mathrm{d}^3 p\, \vec{k} \cdot \frac{\partial f_0}{\partial \vec{p}} \frac{1}{\vec{k} \cdot \vec{v} - \omega} \tag{5.131}$$

which has to vanish for longitudinal electromagnetic waves to propagate. Just as longitudinal electromagnetic waves undergo Landau damping in a plasma, so do potential fluctuations in a stellar-dynamical system.

5.2.4 Jeans length and Jeans mass

Analysing the stability of a self-gravitating system, we need to distinguish potential fluctuations which oscillate from others. Accordingly, we seek the boundary between oscillating and unstable solutions requiring that the frequency should vanish there, $\omega = 0$. If we assume that the unperturbed equilibrium state has a Maxwellian velocity distribution with a velocity dispersion σ^2 and a homogeneous spatial number density n_0,

$$f_0(\vec{v}) = \frac{n_0}{(2\pi\sigma^2)^{3/2}} \mathrm{e}^{-v^2/(2\sigma^2)} , \tag{5.132}$$

the velocity gradient of f_0 required in (5.130) is

$$\frac{\partial f_0}{\partial \vec{v}} = -f_0(\vec{v}) \frac{\vec{v}}{\sigma^2} . \tag{5.133}$$

Without loss of generality, we rotate the coordinate frame such that the positive x axis points into the direction of the wave vector $\vec{k}$, the condition (5.130) simplifies to

$$1 - \frac{4\pi G m n_0}{k^2\sigma^2(2\pi\sigma^2)^{3/2}} \int \mathrm{d}^3 v \frac{k v_x \mathrm{e}^{-v^2/(2\sigma^2)}}{k v_x - \omega} = 0 . \tag{5.134}$$

For $\omega = 0$, the remaining integral is easily solved since

$$\int \mathrm{d}^3 v\, \mathrm{e}^{-v_x^2/(2\sigma^2)} = \left(2\pi\sigma^2\right)^{3/2} , \tag{5.135}$$

and we find from (5.134) the condition

$$k^2\big|_{\omega=0} \equiv k_\mathrm{J}^2 = \frac{4\pi G\rho_0}{\sigma^2} . \tag{5.136}$$

The wave number k_J satisfying this equation is called the *Jeans wave number*. Gravitational instability sets in for larger perturbations, that is, for wave numbers $k < k_\mathrm{J}$ or wave lengths exceeding the so-called Jeans wave length

$$\lambda_\mathrm{J} \equiv \frac{2\pi}{k_\mathrm{J}} = \frac{2\pi\sigma}{\sqrt{4\pi G\rho_0}} = \frac{\sqrt{\pi}\sigma}{\sqrt{G\rho_0}} . \tag{5.137}$$

?

Why is $\omega = 0$ relevant for separating between gravitational stability and instability?

The Jeans wave length or Jeans length defines the Jeans volume λ_J^3 and thereby the *Jeans mass* $M_\mathrm{J} = \rho_0\lambda_\mathrm{J}^3$. An interesting insight follows if we compare the

Jeans mass to the actual mass M of an object and its velocity dispersion. Solving the equation (5.137) for the Jeans length for the density,

$$\rho_0 = \frac{\pi\sigma^2}{G\lambda_\mathrm{J}^2} \;, \tag{5.138}$$

and multiplying with the actual radius R of an object, we find

$$M \approx \rho_0 R^3 = \frac{\pi\sigma^2 R^3}{G\lambda_\mathrm{J}^2} \;. \tag{5.139}$$

However, due to the virial theorem, we must further obey the relation

$$\sigma^2 \approx \frac{GM}{R} \;. \tag{5.140}$$

Eliminating the velocity dispersion between (5.140) and (5.139) shows that

$$R \approx \frac{\lambda_\mathrm{J}}{\sqrt{\pi}} \;. \tag{5.141}$$

The radius of the system is thus necessarily comparable to the Jeans length. This means that the assumption of homogeneity on the scale of the Jeans length cannot be satisfied and that the nature of the instability needs to be studied for each system in detail once its geometry is specified. Nonetheless, the Jeans length defines an order-of-magnitude estimate for the boundary between stability and instability.

5.2.5 Disk potentials

Disk-like structures are of particular importance for stellar-dynamical systems. Imagine a disk in the x-y plane, centred on the coordinate origin, with a spatial matter density

$$\rho(x, y, z) = \Sigma(x, y)\delta_\mathrm{D}(z) \;. \tag{5.142}$$

The disk is thus infinitely thin and has a surface-mass density $\Sigma(x, y)$. Assume further that the disk is axially symmetric about the $\hat{e}_z$ axis. The surface-mass density can then only depend on the radius s in the x-y plane, $\Sigma = \Sigma(s)$. In the appropriate cylindrical coordinates, Poisson's equation then reads

$$\frac{1}{s}\partial_s\left(s\partial_s\phi\right) + \partial_z^2\phi = 0 \tag{5.143}$$

everywhere outside the x-y plane. The structure of this equation suggests a separation ansatz for ϕ,

$$\phi(s, z) = \psi(s)\chi(z) \;. \tag{5.144}$$

Inserting this into (5.143), we obtain

$$\frac{1}{s\psi(s)}\partial_s\left(s\partial_s\psi\right) = \frac{\partial_z^2\chi(z)}{\chi(z)} \;. \tag{5.145}$$

Since the left- and right-hand sides of this equation depend on different variables, s and z, respectively, they must individually equal the same constant, which we call $-k^2$. From the oscillator equation

$$\frac{\partial_z^2\chi(z)}{\chi(z)} = -k^2 \tag{5.146}$$

with negative squared frequency $-k^2$, we immediately infer that $\chi(z)$ must be an exponential function,

$$\chi(z) = \chi_0 \mathrm{e}^{\pm kz} . \tag{5.147}$$

Since the potential should tend to zero far away from the disk, the positive sign applies to negative z and vice versa. The potential therefore decreases exponentially in the direction perpendicular to the disk. Turning to the s dependence, $\psi(s)$ must satisfy the equation

$$\partial_s (s\partial_s\psi) + k^2 s\psi = 0 . \tag{5.148}$$

Substituting $x = ks$ turns this equation into Bessel's differential equation of order zero,

$$x\psi''(x) + \psi'(x) + x\psi(x) = 0 , \tag{5.149}$$

where the prime denotes the derivative with respect to x. Its solution with the appropriate boundary condition that $\psi(s)$ remains regular at the centre $s = 0$ is the ordinary, zeroth-order Bessel function $J_0(x)$. Our solution of Poisson's equation is thus

$$\phi_k(s, z) = \mathrm{e}^{-k|z|} J_0(ks) . \tag{5.150}$$

Any linear superposition of such potential modes $\phi_k(s, z)$ will also be solutions of Poisson's equation.

?

Strictly speaking, we should write the general solution of (5.146) as $\chi(z) = \chi_0 \mathrm{e}^{kz} + \chi_1 \mathrm{e}^{-kz}$ with two constants $\chi_{0,1}$. Why is it appropriate to proceed as described in the text?

For taking the disk into account, we enclose an arbitrary point $(s, 0)$ on the disk by a small cylinder of height h and cross section A such that the disk plane is perpendicular to the symmetry axis of the cylinder and cuts through its centre. We then apply Gauss' law to $\vec{\nabla}^2\phi$ in the cylinder and let the height h become arbitrarily small to find

$$\int \mathrm{d}V \vec{\nabla}^2\phi = \int_{\partial V} \vec{\nabla}\phi \cdot \mathrm{d}\vec{A} \to A \left(\partial_z\phi\big|_{z\to 0+} - \partial_z\phi\big|_{z\to 0-} \right) . \tag{5.151}$$

By Poisson's equation, this integral must equal $4\pi G$ times the mass contained in the cylinder, which is $\Sigma(s)A$. Hence, the surface density causes the discontinuity

$$4\pi G\Sigma(s) = \partial_z\phi\big|_{z\to 0+} - \partial_z\phi\big|_{z\to 0-} \tag{5.152}$$

in the potential gradient $\partial_z\phi$ perpendicular to the disk, Inserting (5.150) here, we see that

$$kJ_0(ks) = 2\pi G\Sigma_k(s) \tag{5.153}$$

must hold for Poisson's equation to be satisfied. However, we generally wish to determine the gravitational potential of disks with arbitrary surface densities $\Sigma(s)$. We can do so if we can find a function $S(k)$ such that

$$\Sigma(s) = \int_0^\infty \mathrm{d}k\, S(k)\Sigma_k(s) = \frac{1}{G} \int_0^\infty \frac{k\mathrm{d}k}{2\pi} S(k) J_0(ks) . \tag{5.154}$$

In this way, the arbitrary surface density $\Sigma(s)$ would then be assembled by linear superposition of modes $\Sigma_k(s)$ that individually satisfy Poisson's equation, and the complete potential would be given by

$$\phi(s, z) = \int_0^\infty \mathrm{d}k\, S(k)\phi_k(s, z) = \int_0^\infty \mathrm{d}k\, \mathrm{e}^{-k|z|} S(k) J_0(ks) . \tag{5.155}$$

To see how such a function $S(k)$ could be found, consider the inverse Fourier transform of an arbitrary, two-dimensional and axi-symmetric function $\hat{f}(k)$,

$$f(s) = \int \frac{k\mathrm{d}k\mathrm{d}\varphi}{(2\pi)^2}\,\hat{f}(k)\,\mathrm{e}^{\mathrm{i}ks\cos\varphi}\ . \qquad (5.156)$$

The azimuthal integral can be carried out, giving

$$\int_0^{2\pi}\mathrm{d}\varphi\,\mathrm{e}^{\mathrm{i}ks\cos\varphi} = \int_0^{2\pi}\mathrm{d}\varphi\,\cos(ks\cos\varphi) = 2\pi\,J_0(ks)\ . \qquad (5.157)$$

?

Why does the solution (5.157) have no imaginary part?

Comparing with (5.154), we see that the sought function $S(k)$ is simply G times the Fourier transform of the surface-density $\Sigma(s)$,

$$S(k) = G\int s\mathrm{d}s\mathrm{d}\varphi\,\Sigma(s)\,\mathrm{e}^{\mathrm{i}ks\cos\varphi} = 2\pi G\int_0^\infty s\mathrm{d}s\,\Sigma(s)J_0(ks)\ . \qquad (5.158)$$

Inserting this back into (5.155) shows that the potential is related to the surface mass density by

$$\phi(s,z) = 2\pi G\int_0^\infty \mathrm{d}k\,\mathrm{e}^{-k|z|}J_0(ks)\int_0^\infty s'\mathrm{d}s'\Sigma(s')J_0(ks')\ . \qquad (5.159)$$

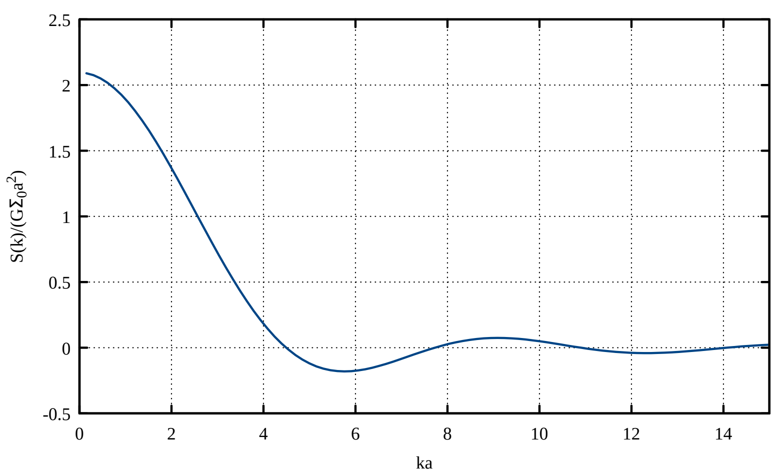

Figure 5.6 The function $S(k)$ is shown here for the Maclaurin disk model, normalised by $G\Sigma_0 a^2$.

5.2.6 Fluid equations for two-dimensional systems

As the simplest example for a rotating disk, we now consider an infinitely thin disk which is rigidly rotating about the z axis with the constant angular velocity $\vec{\Omega} = \Omega\hat{e}_z$. The disk thus fills the x-y plane and has a constant surface-mass density Σ.

We consider perturbations in the plane of the disk and neglect warps or twists. Furthermore, we transform into a co-rotating coordinate frame and study the

Example: The Maclaurin disk

The integrals in (5.159) can be solved analytically only for a surprisingly small class of surface densities $\Sigma(s)$. One example is the so-called *Maclaurin disk*, for which

$$\Sigma(s) = \Sigma_0 \left(1 - \frac{s^2}{a^2}\right)^{1/2} \tag{5.160}$$

for $s \leq a$ and zero otherwise. For this disk model, the function $S(k)$ is

$$S(k) = G\Sigma_0 \sqrt{\frac{2\pi^3 a}{k^3}} J_{3/2}(ka) \;, \tag{5.161}$$

where $J_{3/2}(x)$ is the cylindrical Bessel function of the first kind of fractional order $3/2$. This function is shown in Fig. 5.6. For $s \leq a$, the potential in the plane of the disk is

$$\phi(s, 0) =: \phi_0(s) = \frac{\pi^2 G\Sigma_0}{4a}\left(2a^2 - s^2\right) = -\frac{\pi^2 G\Sigma_0}{4a} s^2 + \text{const} \;. \tag{5.162}$$

Except for an irrelevant constant, the potential within the disk is therefore quadratic in the two-dimensional radius s,

$$\phi_0(s) = -\frac{1}{2}\Omega_0^2 s^2 \;, \quad \Omega_0^2 = \frac{\pi^2 G\Sigma_0}{2a} \;. \tag{5.163}$$

Deriving these results, we have used the integrals

$$\int_0^1 \mathrm{d}x\, x \sqrt{1 - x^2}\, J_0(kx) = \sqrt{\frac{\pi}{2k^3}} J_{3/2}(k) \tag{5.164}$$

and

$$\int_0^\infty \mathrm{d}k\, J_0(ks) J_{3/2}(ka) k^{-3/2} = \frac{1}{4}\sqrt{\frac{\pi}{2a^3}}\left(2a^2 - s^2\right) \;, \tag{5.165}$$

which may not be obvious. ◂

disk in the substantially simpler fluid approximation. We begin with the continuity equation, insert the spatial density

$$\rho(\vec{x}) = \Sigma(t)\delta_{\rm D}(z) \tag{5.166}$$

there and integrate over dz. The result is

$$\partial_t \Sigma + \vec{\nabla} \cdot (\Sigma \vec{v}) = 0 \;, \tag{5.167}$$

where the divergence is now two-dimensional and operates in the x-y plane only. Next, we take the x and y components of the Euler equation of ideal hydrodynamics in the form

$$\Sigma\delta_{\rm D}(z)\frac{{\rm d}\vec{v}}{{\rm d}t} = -\vec{\nabla}P - \Sigma\delta_{\rm D}(z)\vec{\nabla}\phi \;, \tag{5.168}$$

integrate again over dz and obtain

$$\partial_t \vec{v} + \left(\vec{v} \cdot \vec{\nabla}\right)\vec{v} = -\frac{\vec{\nabla}P}{\Sigma} - \vec{\nabla}\phi \;, \tag{5.169}$$

where $\vec{\nabla}$ is again reduced to two dimensions. Poisson's equation reads

$$\vec{\nabla}^2\phi = 4\pi G\Sigma\delta_{\rm D}(z) \tag{5.170}$$

with the three-dimensional Laplacian. We now move into a coordinate system co-rotating with the disk. Since we are then in a non-inertial frame, we must augment Euler's equation with the specific Coriolis and centrifugal force terms, $-2\vec{\Omega} \times \vec{v}$ and $\vec{\Omega}^2\vec{r}$, respectively. With these terms, Euler's equation becomes

$$\frac{\partial\vec{v}}{\partial t} + \left(\vec{v} \cdot \vec{\nabla}\right)\vec{v} = -\frac{\vec{\nabla}P}{\Sigma} - \vec{\nabla}\phi - 2\vec{\Omega} \times \vec{v} + \vec{\Omega}^2\vec{r} \;. \tag{5.171}$$

The physical quantities occuring here have two spatial dimensions, $\vec{v}(x, y, t)$, $\Sigma(x, y, t)$ and so on. For the pressure, we assume a barotropic equation-of-state, $P = P(\Sigma)$.

The unperturbed quantities are obviously a vanishing velocity $\vec{v} = 0$ in the co-rotating frame, a constant surface-mass density $\Sigma = \Sigma_0$, and the pressure $P_0 = P(\Sigma_0)$ according to the equation-of-state. This solution satisfies the continuity equation trivially. Since the pressure gradient vanishes because Σ_0 is constant, Euler's equation reads

$$\vec{\nabla}\phi_0 = \Omega^2\vec{r} \;, \tag{5.172}$$

while Poisson's equation is

$$\vec{\nabla}^2\phi_0 = 4\pi G\Sigma_0\delta_{\rm D}(z) \;. \tag{5.173}$$

Since no direction can be preferred in a homogeneous disk, $\vec{\nabla}\phi_0$ must point along the z axis, which contradicts Euler's equation. Thus, there is no gravitational force yet to balance the centrifugal force. Therefore, we have to assume that the disk can only exist if it is embedded into a surrounding gravitational field which compensates the centrifugal force, such as the halo of a galaxy.

5.2.7 Dispersion relation

As usual in linear stability analyses, we perturb our infinitely extended, rigidly rotating disk by small amounts of surface density $\delta\Sigma$, velocity $\delta\vec{v}$ and gravitational potential $\delta\phi$ and linearise the equations in these perturbations. This implies

$$\begin{aligned} \partial_t\delta\Sigma + \Sigma_0\vec{\nabla}\cdot\delta\vec{v} &= 0\,, \\ \partial_t\delta\vec{v} &= -\frac{c_\mathrm{s}^2}{\Sigma_0}\vec{\nabla}\delta\Sigma - \vec{\nabla}\delta\phi - 2\vec{\Omega}\times\delta\vec{v}\,, \\ \vec{\nabla}^2\delta\phi &= 4\pi G\delta\Sigma\delta_\mathrm{D}(z)\,, \end{aligned} \tag{5.174}$$

where the sound velocity was introduced as the derivative of the pressure with respect to the surface-mass density,

$$c_\mathrm{s}^2 = \left.\left(\frac{\mathrm{d}P(\Sigma)}{\mathrm{d}\Sigma}\right)\right|_{\Sigma_0}\,, \tag{5.175}$$

taken at the unperturbed surface-mass density Σ_0. Next, we decompose the perturbations into plane waves with amplitudes $\delta\Sigma_A$, $\delta\vec{v}_A$ and $\delta\phi_A$,

$$\begin{pmatrix} \delta\Sigma \\ \delta v_x \\ \delta v_y \\ \delta\phi \end{pmatrix} = \begin{pmatrix} \delta\Sigma_A \\ \delta v_{Ax} \\ \delta v_{Ay} \\ \delta\phi_A \end{pmatrix} \mathrm{e}^{\mathrm{i}\left(\vec{k}\cdot\vec{x}-\omega t\right)} \tag{5.176}$$

confined to the x-y plane and therefore valid at $z = 0$. Without loss of generality, we turn the x axis into the direction of the wave vector $\vec{k}$. The first two perturbation equations (5.174) turn into

$$\begin{aligned} \omega\delta\Sigma_A - k\Sigma_0\delta v_{Ax} &= 0\,, \\ \omega\delta v_{Ax} - \frac{c_\mathrm{s}^2 k}{\Sigma_0}\delta\Sigma_A - k\delta\phi_A - 2\mathrm{i}\Omega\delta v_{Ay} &= 0\,, \\ \omega\delta v_{Ay} + 2\mathrm{i}\Omega\delta v_{Ax} &= 0\,, \end{aligned} \tag{5.177}$$

while Poisson's equation needs a more detailed treatment. Within the disk plane, the potential perturbation behaves like the plane wave given by the third equation in (5.176), while the Laplace equation

$$\vec{\nabla}^2\delta\phi = 0 \tag{5.178}$$

must hold otherwise. A separation ansatz demonstrates that this can only be achieved by the function

$$\delta\phi = \delta\phi_\mathrm{A}\mathrm{e}^{\mathrm{i}(kx-\omega t)-|kz|}\,, \tag{5.179}$$

where the modulus is taken of the product kz since $k = k_x$ can have either sign.

?

Verify that the function $\delta\phi$ from (5.179) is the only one satisfying all conditions required here.

The preceding discussion leading to (5.151) and (5.152), specified to our perturbed disk, shows that the potential derivatives in z direction above and below the disk need to obey

$$\left.\partial_z\delta\phi\right|_{z\to0+} - \left.\partial_z\delta\phi\right|_{z\to0-} = 4\pi G\delta\Sigma = 4\pi G\delta\Sigma_A\mathrm{e}^{\mathrm{i}(kx-\omega t)}\,. \tag{5.180}$$

However, we see at the same time from (5.179) that

$$\partial_z\phi\big|_{z\to 0+} - \partial_z\phi\big|_{z\to 0-} = -2|k|\delta\phi_{\mathrm{A}}\mathrm{e}^{\mathrm{i}(kx-\omega t)}\;, \tag{5.181}$$

which implies that the fluctuation amplitudes in the gravitational potential and in the surface-mass density must be related through

$$-\,2|k|\delta\phi_{\mathrm{A}} = 4\pi G\delta\Sigma_{\mathrm{A}}\;. \tag{5.182}$$

This enables us to eliminate the amplitude $\delta\phi_A$ of the potential fluctuations from (5.177), leaving us with the linear system of equations

$$\begin{pmatrix} \omega & -k\Sigma_0 & 0 \\ k\left(\frac{2\pi G}{|k|} - \frac{c_{\mathrm{s}}^2}{\Sigma_0}\right) & \omega & -2\mathrm{i}\Omega \\ 0 & 2\mathrm{i}\Omega & \omega \end{pmatrix} \begin{pmatrix} \delta\Sigma_A \\ \delta v_{Ax} \\ \delta v_{Ay} \end{pmatrix} = 0 \tag{5.183}$$

?

Confirm the dispersion relation (5.184) by setting the determinant of the coefficient matrix in (5.183) to zero.

for the remaining variables $\delta\Sigma_A$, δv_{Ax} and δv_{Ay}.

Non-trivial solutions exist if and only if the determinant of the coefficient matrix vanishes, which leads us to the dispersion relation

$$\omega\left(\omega^2 - 4\Omega^2\right) + \omega k^2\Sigma_0\left(\frac{2\pi G}{|k|} - \frac{c_{\mathrm{s}}^2}{\Sigma_0}\right) = 0 \tag{5.184}$$

for the perturbations of the disk. This shows that perturbations must either be stationary, $\omega = 0$, or obey

$$\omega^2 = 4\Omega^2 + k^2c_{\mathrm{s}}^2 - 2\pi G|k|\Sigma_0\;. \tag{5.185}$$

This dispersion relation describes the non-stationary, propagating modes of the perturbed, rigidly rotating, uniform disk. The modes are stable for $\omega^2 \geq 0$ and unstable for $\omega^2 < 0$.

5.2.8 Toomre's criterion

Let us now analyse the dispersion relation (5.185). If $\Omega = 0$, which is certainly not the most exciting case of a rotating disk, the disk is unstable if the wave number satisfies

$$|k| < k_{\mathrm{J}} \equiv \frac{2\pi G\Sigma_0}{c_{\mathrm{s}}^2}\;, \tag{5.186}$$

where k_{J} plays the rôle of the Jeans wave number for the disk. If the sound speed can be arbitrarily low, $c_{\mathrm{s}} \to 0$, perturbations are unstable for arbitrarily large k or arbitrarily small wave length. The rate of the exponential growth of unstable perturbations is given by the imaginary part of ω,

$$\gamma = \mathrm{Im}\omega = \left(2\pi G\Sigma_0|k| - k^2c_{\mathrm{s}}^2\right)^{1/2}\;. \tag{5.187}$$

For a cold disk, $c_{\mathrm{s}} \to 0$, small perturbations with $\lambda \to 0$ and $|k| \to \infty$ grow at a rate increasing linearly with $|k|$, i.e. cold, non-rotating disks fragment violently on small scales.

This violent fragmentation is not suppressed by rotation either. For $c_s \to 0$, the oscillation frequency ω of the linear perturbations becomes imaginary for wave numbers

$$|k| > \frac{2\Omega^2}{\pi G \Sigma_0} \ , \tag{5.188}$$

i.e. even then the instability sets in on the smallest scales.

Pressure and rotation are therefore not able to stabilise the disk individually. For $\Omega = 0$, (5.186) shows that warm disks are unstable on large scales,

$$|k| < \frac{2\pi G \Sigma_0}{c_s^2} \ , \tag{5.189}$$

while cold disks with vanishing pressure are unstable on small scales despite any rotation, as we have seen in (5.188). However, pressure and rotation can be stabilising if they act *together*, since then the dispersion relation (5.185) has a minimum where

$$0 = \frac{\partial \omega^2}{\partial k} = \frac{\partial}{\partial k}\left(4\Omega^2 + k^2 c_s^2 - 2\pi G |k| \Sigma_0\right) \ , \tag{5.190}$$

which yields

$$2|k|c_s^2 = 2\pi G \Sigma_0 \quad \text{or} \quad |k| = \frac{\pi G \Sigma_0}{c_s^2} = \frac{k_J}{2} \ . \tag{5.191}$$

The disk can be stable if and only if $\omega^2 \geq 0$ at this wave number because it is then positive for all wave numbers. Thus, the condition for global stability is $\omega^2(k_J/2) \geq 0$ or

$$4\Omega^2 - \left(\frac{\pi G \Sigma_0}{c_s^2}\right) \geq 0 \ , \tag{5.192}$$

which constrains the sound speed c_s and the angular velocity Ω for a globally stable disk with mean surface-mass density Σ_0 by *Toomre's criterion*

$$\frac{c_s \Omega}{G \Sigma_0} \geq \frac{\pi}{2} \approx 1.57 \ . \tag{5.193}$$

A similar criterion can also be derived for collision-less systems (recall that we had adopted the fluid approximation!). Then,

$$\frac{c_s \Omega}{G \Sigma_0} \gtrsim 1.68 \tag{5.194}$$

is the condition for global stability.

Problems

1. Show that the scalar virial theorem reads

$$2K = -pU \tag{5.195}$$

if the potential energy scales like $U \propto r^{-p}$ with the separation r between two particles. Derive for which values of p the heat capacity of self-gravitating systems is negative.

5.3 Dynamical Friction

This final section discusses the friction experienced by a test mass moving through an infinite system of other masses due to gravity. We first calculate the velocity changes (5.213) and (5.215) of the test mass perpendicular and parallel to its direction of motion due to encounters with the surrounding masses and average over impact parameters to arrive at Chandrasekhar's formula (5.228) for the acceleration due to dynamical friction. A specialisation of this formula for a Maxwellian velocity distribution in the surrounding system of masses is given in (5.232).

5.3.1 Deflection of point masses

An interesting effect occurs if a mass M moves through a system of masses m which are homogeneously distributed around the mass M. Although the motion of the masses can be considered collision-less, a deceleration occurs which is called *dynamical friction*.

Let us begin analysing a single two-body encounter between the mass M and a mass m, with $\vec{v}_M$ and $\vec{v}_m$ being their respective velocities and $\vec{x}_M$ and $\vec{x}_m$ being their positions. We introduce the separation vector

$$\vec{r} \equiv \vec{x}_m - \vec{x}_M \tag{5.196}$$

from M to m and the relative velocity

$$\vec{v} \equiv \dot{\vec{r}} = \vec{v}_m - \vec{v}_M \tag{5.197}$$

of m with respect to M. The two-body system of two point masses obeys an effective equation of motion around a fixed force centre of a single body with the *reduced mass*,

$$\left(\frac{mM}{m+M}\right)\ddot{\vec{r}} = -\frac{GMm}{r^2}\hat{e}_r \equiv -\frac{\alpha}{r^2}\hat{e}_r \;, \tag{5.198}$$

where $\hat{e}_r$ is the unit vector in radial direction away from M. By definition of the centre-of-mass $\vec{X}$,

$$\vec{X} = \frac{m\vec{x}_m + M\vec{x}_M}{m+M} \;, \tag{5.199}$$

its velocity remains unchanged,

$$\dot{\vec{X}} = \frac{m\vec{v}_m + M\vec{v}_M}{m+M} = 0 \;. \tag{5.200}$$

Any changes $\Delta\vec{v}_m$ and $\Delta\vec{v}_M$ in the velocites of m and M are thus related by

$$M\Delta\vec{v}_M = -m\Delta\vec{v}_m \tag{5.201}$$

with each other and by

$$\Delta\vec{v}_M = -\frac{m}{M}\Delta\vec{v}_m = -\frac{m}{M}(\Delta\vec{v} + \Delta\vec{v}_M) \quad \text{or} \quad \Delta\vec{v}_M = -\frac{m}{M+m}\,\Delta\vec{v} \tag{5.202}$$

to any change

$$\Delta\vec{v} = \Delta\vec{v}_m - \Delta\vec{v}_M \tag{5.203}$$

in the relative velocity of m and M. We shall now determine the relative velocity change $\Delta\vec{v}$ in a two-body encounter.

The fictitous particle with the reduced mass, $Mm/(M+m)$, follows a hyperbolic orbit around the (resting) centre of force.

Earlier in this book, we came across another situation where one particle moves on a hyperbolic orbit around another one, namely when we studied the emission of a plasma electron scattering off an ion. We can thus refer to the treatment there. We recall the relation (2.99) between the distance r of the orbiting particle from the force centre and the polar angle φ, from which we read off that the particle reaches infinity if and when

$$\cos\varphi = -\frac{1}{\varepsilon}\,, \tag{5.204}$$

where ε is the orbit's eccentricity. From (2.100), we infer that the squared eccentricity is

$$\varepsilon^2 = 1 + \frac{2El^2}{\alpha^2\mu}\,, \tag{5.205}$$

where $\alpha = GMm$ replaces the product Ze^2 in (2.100) since now the coupling is gravitational rather than electromagnetic, and the reduced mass μ replaces the electron mass m. Let now v be the velocity at infinity of our fictituous particle with the reduced mass μ, and b its impact parameter. Then, its conserved angular momentum is $l = \mu vb$, and its equally conserved energy is $E = \mu v^2/2$. We can then re-write the squared eccentricity as

$$\varepsilon^2 = 1 + \frac{\mu^2b^2v^4}{\alpha^2}\,. \tag{5.206}$$

Since the total scattering angle is $\theta = 2\varphi - \pi$, we have

$$\sin\frac{\theta}{2} = \sin\left(\varphi - \frac{\pi}{2}\right) = -\cos\varphi = \frac{1}{\varepsilon}\,. \tag{5.207}$$

Since the cosine of $\theta/2$ is

$$\cos\frac{\theta}{2} = \sqrt{1-\sin^2\frac{\theta}{2}} = \frac{\sqrt{\varepsilon^2-1}}{\varepsilon}\,, \tag{5.208}$$

we find for the scattering angle itself

$$\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} = 2\frac{\sqrt{\varepsilon^2-1}}{\varepsilon^2} = \frac{2\mu bv^2\alpha}{\mu^2b^2v^4+\alpha^2} \tag{5.209}$$

and

$$\cos\theta = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} = \frac{\varepsilon^2-2}{\varepsilon^2} = \frac{\mu^2b^2v^4-\alpha^2}{\mu^2b^2v^4+\alpha^2}\,. \tag{5.210}$$

5.3.2 Velocity changes

The fictitous particle is approaching from infinity with velocity v. When it is leaving towards infinity, it will have been scattered by the angle θ. Relative to its incoming direction, its outgoing velocity will have the perpendicular and parallel components

$$v_\perp = -v\sin\theta = -\frac{2\mu b v^3\alpha}{\mu^2 b^2 v^4 + \alpha^2} \tag{5.211}$$

and

$$v_\parallel = v\cos\theta = v\,\frac{\mu^2 b^2 v^4 - \alpha^2}{\mu^2 b^2 v^4 + \alpha^2}\ , \tag{5.212}$$

where (5.209) and (5.210) were used. Since the velocity change perpendicular to the incoming direction is $\Delta v_\perp = v_\perp$, we find from (5.202) that the velocity change of the mass M perpendicular to its initial direction of motion is

$$\Delta v_{M\perp} = -\frac{m}{M+m}v_\perp = -\frac{\mu}{M}v_\perp = \frac{2\mu^2 b v^3\alpha}{M\left(\mu^2 b^2 v^4 + \alpha^2\right)}\ . \tag{5.213}$$

Parallel to the incoming direction, we have

$$\Delta v_\parallel = v - v_\parallel = -\frac{2v\alpha^2}{\mu^2 b^2 v^4 + \alpha^2}\ , \tag{5.214}$$

hence the velocity change of the mass M in the direction parallel to its initial motion is

$$\Delta v_{M\parallel} = -\frac{\mu}{M}\Delta v_\parallel = \frac{2v\mu\alpha^2}{M\left(\mu^2 b^2 v^4 + \alpha^2\right)}\ . \tag{5.215}$$

5.3.3 Chandrasekhar's formula

Having studied the effect of a single encounter on the velocity of the mass M, we shall proceed to determine the combined effect of many encounters. If the mass M is moving through a homogeneous "sea" of masses m, all velocity changes (5.213) perpendicular to the direction of motion must cancel, while the parallel velocity changes (5.215) must add up. Therefore, the mass M will experience a steady deceleration parallel to its direction of motion from the combined effect of many encounters with the masses m. Let $f(\vec{v}_m)$ be the phase-space density of the stars with mass m which constitute that background "sea" of point masses. Then, the rate at which the mass M encounters collisions with stars with an impact parameter between b and $b + \mathrm{d}b$ is

$$(2\pi b\mathrm{d}b)\cdot v\cdot\left(f(\vec{v}_m)\mathrm{d}^3 v_m\right)\ , \tag{5.216}$$

where the first factor is the area of the ring with radius b and width $\mathrm{d}b$ and the third is the spatial number density of masses m with velocity v_m. The velocity v appearing in between is the relative velocity of M and m. According to (5.215), these collisions change the velocity of M by

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = \vec{v}f(\vec{v}_m)\mathrm{d}^3 v_m\int_0^{b_{\max}} 2\pi b\mathrm{d}b\,\frac{2v\mu\alpha^2}{M\left(\mu^2 b^2 v^4 + \alpha^2\right)}\ , \tag{5.217}$$

The integral is easily carried out and returns a logarithm,

$$\begin{aligned}\int_0^{b_{\max}} 2\pi b\mathrm{d}b\,\frac{2v\mu\alpha^2}{M\left(\mu^2 b^2 v^4 + \alpha^2\right)} &= \frac{2\pi v\mu\alpha^2}{M}\int_0^{b_{\max}^2}\frac{\mathrm{d}(b^2)}{\mu^2 b^2 v^4 + \alpha^2}\\ &= \frac{2\pi\alpha^2}{M\mu v^3}\ln\left(1+\Lambda^2\right)\ ,\end{aligned} \tag{5.218}$$

where the abbreviation

$$\Lambda := \frac{b_{\max}\mu v^2}{\alpha} = \frac{b_{\max}v^2}{G(M+m)} \tag{5.219}$$

was introduced. Since $\vec{v}$ in (5.217) is the relative velocity between the masses m and M, we need to set $\vec{v} = \vec{v}_m - \vec{v}_M$. Inserting this expression and (5.218) into (5.217), we find

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = \frac{2\pi\alpha^2}{M\mu}\ln\left(1+\Lambda^2\right)\frac{\vec{v}_m - \vec{v}_M}{\left|\vec{v}_m - \vec{v}_M\right|^3} f(\vec{v}_m)\,\mathrm{d}^3 v_m \,. \tag{5.220}$$

The quantity Λ is typically $\Lambda \gg 1$, allowing us to approximate

$$\ln(1+\Lambda^2) \approx \ln\Lambda^2 = 2\ln\Lambda \,. \tag{5.221}$$

Typical values for this so-called *Coulomb logarithm* are

$$5 \lesssim \ln\Lambda \lesssim 20 \,. \tag{5.222}$$

The expression (5.220) is the deceleration of the mass M by those of the surrounding stars with mass m whose velocity falls within the volume element $\mathrm{d}^3 v_m$ around $\vec{v}_m$ in velocity space. We obtain the total deceleration only after a further integration over all velocities $\vec{v}_m$. We proceed to doing so in two steps. First, we specialise to isotropic velocity distributions, $f(\vec{v}_m) = f(v_m)$. We further abbreviate $\vec{x} = \vec{v}_m$ and $\vec{y} = \vec{v}_M$ and begin by simplifying the velocity integral as

$$\int \mathrm{d}^3x\,\frac{f(x)\,(\vec{x}-\vec{y})}{\left|\vec{x}-\vec{y}\right|^3} = -\int \mathrm{d}^3x\,f(x)\vec{\nabla}\frac{1}{\left|\vec{x}-\vec{y}\right|} = \vec{\nabla}_{\vec{y}}\int \mathrm{d}^3x\,\frac{f(x)}{\left|\vec{x}-\vec{y}\right|} \,. \tag{5.223}$$

For solving the remaining integral, we turn the coordinate system such that $\vec{y}$ points into the $\hat{e}_z$ direction, introduce spherical polar coordinates and continue writing

$$\int \mathrm{d}^3x\,\frac{f(x)}{\left|\vec{x}-\vec{y}\right|} = 2\pi\int_0^\infty x^2\mathrm{d}x\,f(x)\int_{-1}^1\frac{\mathrm{d}\mu}{\sqrt{x^2+y^2-2xy\mu}} \,, \tag{5.224}$$

where the direction cosine $\mu = \cos\theta$ was introduced. The factor of 2π out front is the result of the azimuthal integration. The μ integral is now easily carried out, giving

$$\int_{-1}^1\frac{\mathrm{d}\mu}{\sqrt{x^2+y^2-2xy\mu}} = \frac{|x+y|-|x-y|}{xy} = \begin{cases} \frac{2}{x} & (x>y) \\ \frac{2}{y} & (\text{else}) \end{cases} \,. \tag{5.225}$$

?

Carry out the integral in (5.225) yourself.

Returning to (5.224), we write

$$\int \mathrm{d}^3x\,\frac{f(x)}{\left|\vec{x}-\vec{y}\right|} = 4\pi\left(\frac{1}{y}\int_0^y x^2\mathrm{d}x\,f(x) + \int_y^\infty x\mathrm{d}x\,f(x)\right) \,. \tag{5.226}$$

The gradient with respect to $\vec{y}$, required by (5.223), finally leaves us with

$$\int \mathrm{d}^3x\,\frac{f(x)\,(\vec{x}-\vec{y})}{\left|\vec{x}-\vec{y}\right|^3} = -4\pi\frac{\vec{y}}{y^3}\int_0^y x^2\mathrm{d}x\,f(x) \,. \tag{5.227}$$

Restoring the velocities $\vec{v}_M = \vec{y}$ and $\vec{v}_m = \vec{x}$ in this expression and returning with it to the deceleration (5.220), we obtain *Chandrasekhar's equation* for dynamical friction (Figure 5.7),

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = -16\pi^2 G^2 m(M+m) \ln\Lambda \frac{\vec{v}_M}{v_M^3} \int_0^{v_M} v_m^2 \mathrm{d}v_m f(v_m) \ , \tag{5.228}$$

where we have also expanded $\alpha = GMm$ and $\mu = Mm/(M+m)$.

Limiting cases are instructive. If v_M is small compared to the typical velocity of the stars m, the remaining integral can be approximated by

$$\int_0^{v_M} \mathrm{d}v_m\, v_m^2 f(v_m) \approx \frac{v_M^3}{3} f(0) \ . \tag{5.229}$$

In this case, the dynamical friction becomes proportional to the velocity $\vec{v}_M$,

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = -\frac{16\pi^2}{3} G^2 m(M+m) \ln\Lambda\, f(0)\, \vec{v}_M \ , \tag{5.230}$$

which is characteristic for Stokes-type friction. In the opposite limiting case of sufficiently large v_M, the integral covers most or all of the velocity distribution of the masses m and thus converges to a constant. Then, the friction force becomes proportional to v_M^{-2}.

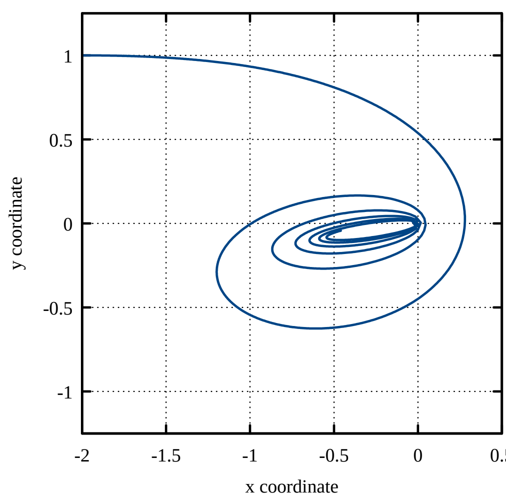

Figure 5.7 An example is shown for the trajectory of a test particle under the combined effect of gravity and dynamical friction.

For a Maxwellian velocity distribution with velocity dispersion σ,

$$f(v_m) = \frac{n_0}{(2\pi\sigma^2)^{3/2}} \mathrm{e}^{-v^2/(2\sigma^2)} \ , \tag{5.231}$$

the friction force becomes

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = -4\pi G^2 \ln\Lambda\, (M+m)\rho_0 \frac{\vec{v}_M}{v_M^3} \left[\mathrm{erf}(X) - \frac{2X}{\sqrt{\pi}} \mathrm{e}^{-X^2} \right] \ , \tag{5.232}$$

where $X \equiv v_M/(\sqrt{2}\sigma)$ is the velocity v_M appropriately scaled by the velocity dispersion. The mass density ρ_0 of the masses m is $\rho_0 = mn_0$. If $M \gg m$, $(M + m) \approx M$, and the friction only depends on the density ρ_0 of the scatteres stars, but not on their mass any more,

$$\frac{\mathrm{d}\vec{v}_M}{\mathrm{d}t} = -4\pi G^2 \ln\Lambda\,\rho_0 M \frac{\vec{v}_M}{v_M^3}\left[\mathrm{erf}(X) - \frac{2X}{\sqrt{\pi}}\mathrm{e}^{-X^2}\right] . \tag{5.233}$$

In this case, the friction force is proportional to M^2 because the deceleration is proportional to M. An example for an orbit decaying via the dynamical friction in a system with Maxwellian velocity distribution is shown in Fig. 5.7.

Problems

1. Approximate the velocity distribution of a hypothetical stellar-dynamical system by the step function

$$f(v_m) = \Theta(v_0 - v_m) \tag{5.234}$$

with some maximal velocity $v_0 > 0$.

   (a) Solve Chandrasekhar's equation for dynamical friction (5.228) for a star moving through this system with an initial velocity $V_{M0} > v_0$.

   (b) How long will it take for the star to reach v_0, and when will it stop completely?

Suggested further reading: [19, 4]

Brief summary and concluding remarks

Our introductory tour through theoretical astrophysics is coming to an end. Its main goal should be seen as achieved if it could reveal the roots of four main branches of the field: the largely classical physics of electromagnetic radiation processes, ideal and viscous hydrodynamics, the physics of plasmas with and without magnetic fields, and stellar dynamics.

Larmor's equation, underlying the strictly classical electromagnetic radiation processes, follows directly from the Liénard-Wiechert potentials and thus from the retarded Greens function of electrodynamics. The back-reaction of the radiation on the radiating charges needs to be added by hand to classical electrodynamics because it is a linear theory. For Compton scattering, the photon picture needs to be introduced, and quantum mechanics, in particular Fermi's Golden Rule, is required to describe the internal degrees of freedom in systems interacting with radiation.

Hydrodynamics, including viscous, relativistic and magnetised fluids, follows from the conservation law for the energy-momentum tensor, which can be derived from moments of the Boltzmann equation. The phenomenology of the resulting equations is very rich, but global statements can be derived by integration. Powerful examples are Bernoulli's law, which is an integral of Euler's equation, and the Rankine-Hugoniot jump conditions at shocks. Linear stability analysis, following a well-defined procedure, reveals a rich variety of dispersion relations and instabilities, some of which were discussed. Turbulence, at the boundary between ordered, macroscopic and unordered, microscopic motion, could only be touched briefly.

Plasma physics adds electromagnetic properties to a fluid which are encapsulated in the dielectric tensor. Magnetic fields add two types of force to a plasma, one due to gradients in the magnetic pressure, the other due to the bending of field lines. Through the induction equation, the fluid flow acts back on the magnetic field. Ambipolar diffusion was introduced as an effect arising from non-ideal coupling between a plasma a neutral fluid, and the battery mechanism was mentioned as an example for how magnetic fields can be generated. The very rich field of magneto-hydrodynamic stability analysis could only be discussed to the level of Alfvén waves and the fast and slow hydromagnetic modes.

Stellar dynamics finally has the same root as hydrodynamics, namely the Liouville or Boltzmann equations, depending on what degree of interaction between particles is to be included. Three main differences to hydrodynamics arise: The long-range gravitational interaction between the particles cannot be shielded, the absence of collisions allows anisotropic particle orbits and thus an anisotropic pressure, and self-gravitating systems have a negative heat capacity and thus no stable equilibrium state. Again, the rich subject of stability analyses of self-gravitating systems could only briefly be touched.

Equally important as recognising the foundations of theoretical astrophysics is becoming aware of their limitations due to idealising assumptions. Some care was taken to clearly specify the assumptions made. If this book can enable its readers to ask and address their own questions on the theory of astrophysical phenomena, it has reached its goal.

Bibliography

[1] B.W. Carroll, D.A. Ostlie, *An Introduction to Modern Astrophysics* (Addison-Wesley, 2006)

[2] J.D. Jackson, *Classical electrodynamics* (John Wiley & Sons, 1998)

[3] L.D. Landau, E.M. Lifshitz, *Course of theoretical physics. II. Classical theory of fields* (Butterworth Heinemann, 1987)

[4] E.M. Lifshitz, L.P. Pitaevskii, *Course of theoretical physics. X. Physical kinetics* (Butterworth Heinemann, 1981)

[5] J.C. Mather, E.S. Cheng, D.A. Cottingham, R.E. Eplee, Jr., D.J. Fixsen, T. Hewagama, R.B. Isaacman, K.A. Jensen, S.S. Meyer, P.D. Noerdlinger, S.M. Read, L.P. Rosen, R.A. Shafer, E.L. Wright, C.L. Bennett, N.W. Boggess, M.G. Hauser, T. Kelsall, S.H. Moseley, Jr., R.F. Silverberg, G.F. Smoot, R. Weiss, D.T. Wilkinson, Astrophysical J. **420**, 439 (1994). DOI 10.1086/173574

[6] F. Shu, *The physics of astrophysics. I. Radiation* (University Science Books, 1991)

[7] G.B. Rybicki, A.P. Lightman, *Radiative processes in astrophysics* (John Wiley & Sons, 1979)

[8] P.A.M. Dirac, *The principles of quantum mechanics* (Oxford University Press, 1988)

[9] J.J. Sakurai, S.F. Tuan, *Modern quantum mechanics* (Prentice Hall, 1993)

[10] S.N. Shore, *Astrophysical Hydrodynamics: An Introduction* (Wiley-VCH, 2007)

[11] N. Straumann, *General relativity with applications to astrophysics* (Springer, 2004)

[12] S. Carroll, *Spacetime and Geometry: Introduction to General Relativity* (Prentice Hall, 2003)

[13] A.R. Choudhuri, *The Physics of Fluids and Plasmas: An Introduction for Astrophysicists* (Cambridge University Press, 1998)

[14] F. Shu, *The physics of astrophysics. II. Gas dynamics* (University Science Books, 1992)

[15] L.D. Landau, E.M. Lifshitz, *Course of theoretical physics. VII. Fluid mechanics* (Butterworth Heinemann, 1987)

[16] S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability* (Dover, 1981)

[17] F. Reif, *Fundamentals of statistical and thermal physics* (McGraw-Hill, 1987)

[18] S. Ishimaru, *Basic principles of plasma physics. A statistical approach* (Addison-Wesley, 1980)

[19] J. Binney, S. Tremaine, *Galactic Dynamics* (Princeton University Press, 1987)

Index

Understanding astronomical objects requires knowledge and methods from different branches of theoretical physics: we diagnose these objects mostly by the light we receive; the observed phenomena often have to do with the flow of fluids, sometimes ionised, sometimes magnetised; and the measured velocities reflect dynamics driven by gravitational fields. Courses in theoretical physics lay the foundation in classical and quantum mechanics, electrodynamics, and thermodynamics, but a gap remains between this foundation and its application to astrophysics. These lecture notes build upon the core courses in theoretical physics and provides the methods for understanding astrophysics theoretically.

About the Author

Matthias Bartelmann is professor for theoretical astrophysics at Heidelberg University. He mostly addresses cosmological questions, concerning in particular the formation and evolution of cosmic structures.

UNIVERSITÄT HEIDELBERG
ZUKUNFT
SEIT 1386