Numerical astrophysics

Lecture notes

Wolfgang Dobler & Michael Stix

Revision: 1.48 , 27th September 2005

Contents

Introduction

In our scientific education, we grow up with many examples of analytically solvable problems; most of them are described by linear differential equations, since only for linear equations a systematic theory for constructing solutions is available. While analytically tractable problems are much more instructive and allow far deeper understanding than other problems, they still represent just a subset of measure zero of the full class of interesting physical problems. For all the rest, we have basically just one choice — to solve them numerically.

Pioneered in the 1950s (and even before, as far as the theoretical basis is concerned) and initially (ab)used for the creation of devices for mass-destruction, *computational physics*, and in particular the numerical solution of partial differential equations (most importantly the equations of hydrodynamics), is now a well-established approach to solve complicated, nonlinear problems, which cannot be successfully addressed by other methods.

This does however not mean that the analytical approach is outdated in any sense. Good numerical modelling is only possible with good analytical skills and physical understanding. Moreover, the usefulness of a numerical code may crucially depend on choosing the right variables and applying the right idealisations in the model — both of which require analytical penetration of the physics involved. After all, this is just what physics itself is all about: finding minimal models of a physical system and describing it using the clearest concepts and most efficient variables.

Scheduled with only one lecture block per week during a short semester, the present lecture course can only scratch the surface of numerical methods and computational physics. This implies that only a few central concepts of numerical mathematics and of the programming language (in our case *IDL*) will be explicitly introduced, and even these will just be discussed in an application-oriented fashion. Those students previously exposed to the extensive volumes of literature on numerical methods might consider this an advantage.

The course will address a range of astrophysical problems, from celestial mechanics over hydrodynamics and magnetohydrodynamics to (some very special questions of) general relativity. While fluid dynamics will play a prominent role, the problems solved will never go beyond one-dimensional problems (i. e. one space dimension, plus time-dependency); a few problems will even be technically

'zero-dimensional', like the N-body problem we will start with, or the Friedman equations, which are both described by ordinary differential equations. One reason for this restriction is that the interactive tools we will use are not the most efficient ones, and dealing with multidimensional problems one would eventually work with compiled languages like *Fortran 90*. Moreover, visualisation and analysis of results becomes increasingly complex if more than one space dimension is involved.

If you really want to learn how to solve multi-dimensional problems, the methods you will acquire in the present course will provide a good starting point. A number of scientifically relevant fluid dynamics codes (some of them including magnetic fields, radiative transfer, or forced turbulence) are based on exactly the methodology we will apply in one space dimension.

The numerical experiments and calculations for this course will be programmed in *IDL* (Interactive Data Language), a proprietary software package that is quite popular in astronomy. This choice has mostly historical reasons and IDL could be replaced by any other tool that (a) provides a full-featured programming language, including basic plotting functions, (b) can be used interactively, and (c) provides *array syntax* to compactly and efficiently manipulate arrays of data.

The best alternative to IDL is probably *PerlDL*, an array extension to the powerful high-level programming language *Perl*.[1] Other options (although at least some of them are far slower on some tasks) are *Matlab*-derivates such as *Octave* or *Scilab*.

The one basic message this course is trying to spread is that, using the right tools, the numerical solution of differential equations in general, and the partial differential equations of hydrodynamics in particular, is in fact something natural and simple.

[1] For users insisting on stricter morphology ('less line-noise'), *Python-numeric* tries to fill a similar gap, although it does not appear to have a considerable user community.

Chapter 1

Numerical Methods I – ordinary differential equations

1.1 Time stepping schemes

Explicit systems of ordinary differential equations (*ODEs*)

$$\frac{dy_i}{dt} = F_i(y_k, t) \tag{1.1}$$

are solved by *time stepping methods* which, given the state vector $y_i(t_0)$ at a given time t_0, yield an approximation to the state vector at time $t_0+\delta t$, evaluating the right-hand side $F(\cdot,\cdot)$ one or several times in the process. Multi-step methods (*predictor-corrector* schemes) additionally use state vectors from a few previous time steps, but we will focus on single-step methods here.

1.2 Euler scheme

The simplest and most transparent (and often also the least efficient) time stepping scheme is the *Euler scheme*. Here the time derivative dy/dt in Equ. (1.1) is approximated by a forward difference

$$\frac{dy}{dt}(t_n) \longmapsto \frac{y(t_n+\delta t) - y(t_n)}{\delta t} \ , \tag{1.2}$$

and the resulting scheme takes the form

$$y_i^{(n+1)} = y_i^{(n)} + \delta t\, F_i\left(y_k^{(n)}, t_n\right) + O\left(\delta t^2\right) \ , \tag{1.3}$$

where $y^{(n)} \equiv y(t_n)$.

The Euler scheme (1.3) is a first-order scheme, i. e. the error when integrating over a finite time T is

$$R = \frac{T}{\delta t} O\left(\delta t^2\right) = O\left(\delta t\right) \ . \tag{1.4}$$

See Fig. 1.1 for an illustration of Euler stepping.

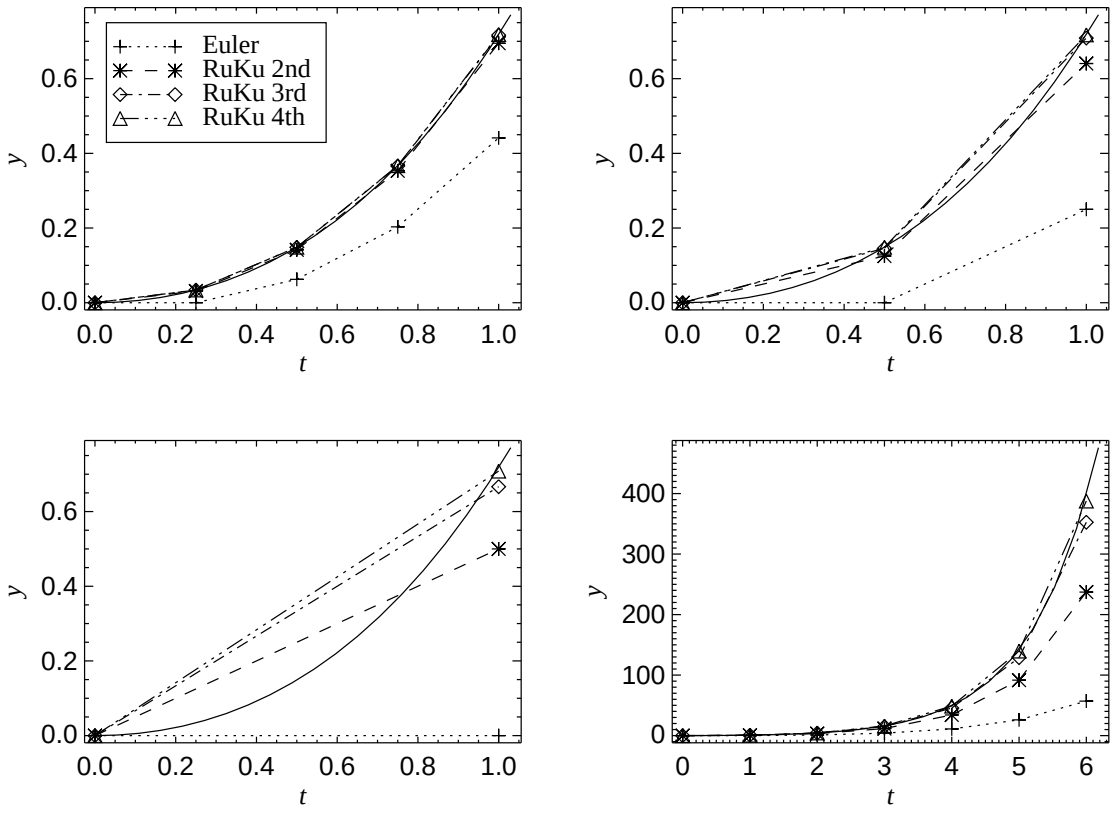

Figure 1.1: Comparison of time stepping schemes of different order for a few steps of Eq. (1.7). Top left: $\delta t = 0.25$; top right: $\delta t = 0.5$; bottom: $\delta t = 1$.

1.3 Runge–Kutta schemes

A much more accurate class of single-step schemes are the *Runge–Kutta schemes*. In these schemes, a first Euler step is consecutively improved by evaluating the right-hand side $F_i(y_k, t)$ several more times for particular arguments y_k and t. There are many Runge–Kutta schemes in use; the best way to familiarise with a given scheme is to pick a simple, not too trivial differential equation and work through a few time steps with a calculator. A convenient sample ODE is

$$\frac{dy}{dt} = y + t \; ; \qquad t_0 = 0 \; , \; y(t_0) = 0 \; , \tag{1.5}$$

which has the exact solution

$$y(t) = e^t - 1 - t \; . \tag{1.6}$$

Below, we give a few schemes, two of which we will use in the future. A comparison of different schemes for two different time steps is shown in Fig. 1.1.

A second-order scheme (mostly for illustration). The most obvious improvement over the Euler scheme is to use the central difference operator

$$\frac{dy}{dt}(t_n + \delta t/2) \longmapsto \frac{y(t_n+\delta t) - y(t_n)}{\delta t} \ , \tag{1.7}$$

which, unlike (1.2), is second-order accurate. But in order to apply this for time stepping, we need some estimate for $\frac{dy}{dt}(t_n + \delta t/2)$. I turns out that a first-order estimate will do here, thus we can use an Euler step of width $\delta t/2$, yielding

$$y(t_0 + \delta t/2) \approx y_1 = y(t_0) + \delta t F(y(t_0), t(0)) \ . \tag{1.8}$$

With this, we then obtain the second-order estimate

$$y(t_0 + \delta t) \approx y(t_0) + \delta t F(y_1, t + \delta t/2) \ . \tag{1.9}$$

A more concise representation of this scheme is given by the following tableau

t	y	
t_0	y_0	$k_1 = \delta t F(y_0, t_0)$
$t_1 = t_0 + \frac{1}{2}\delta t$	$y_1 = y_0 + \frac{1}{2}k_1$	$k_2 = \delta t F(y_1, t_1)$
$t = t_0 + \delta t$	$y = y_0 + k_2 + O\left(\delta t^3\right)$	

The scheme requires two evaluations of the right-hand side per step and is second-order accurate. It is rarely used in practise, as opposed to the third- and fourth-order schemes given below.

A third-order scheme. We will use this scheme for time-stepping partial differential equations later, because it requires less memory than many comparable schemes.

t	y	
t_0	y_0	$k_1 = \delta t F(y_0, t_0)$
$t_1 = t_0 + \frac{8}{15}\delta t$	$y_1 = y_0 + \frac{8}{15}k_1$	$k_2 = \delta t F(y_1, t_1)$
$t_2 = t_0 + \frac{2}{3}\delta t$	$y_2 = y_0 + \frac{1}{4}k_1 + \frac{5}{12}k_2$	$k_3 = \delta t F(y_2, t_2)$
$t = t_0 + \delta t$	$y = y_0 + \frac{1}{4}k_1 + 0 \cdot k_2 + \frac{3}{4}k_3 + O\left(\delta t^4\right)$	

When integrating over a finite time, the error term will be

$$R = \frac{T}{\delta t} O\left(\delta t^4\right) = O\left(\delta t^3\right) \ . \tag{1.10}$$

The "classical' fourth-order Runge–Kutta scheme. This scheme is very popular for ODEs and systems of ODEs. For PDEs, it may well deliver more accuracy than is compatible with the space discretisation, so it may be better to use a 3rd-order scheme only.

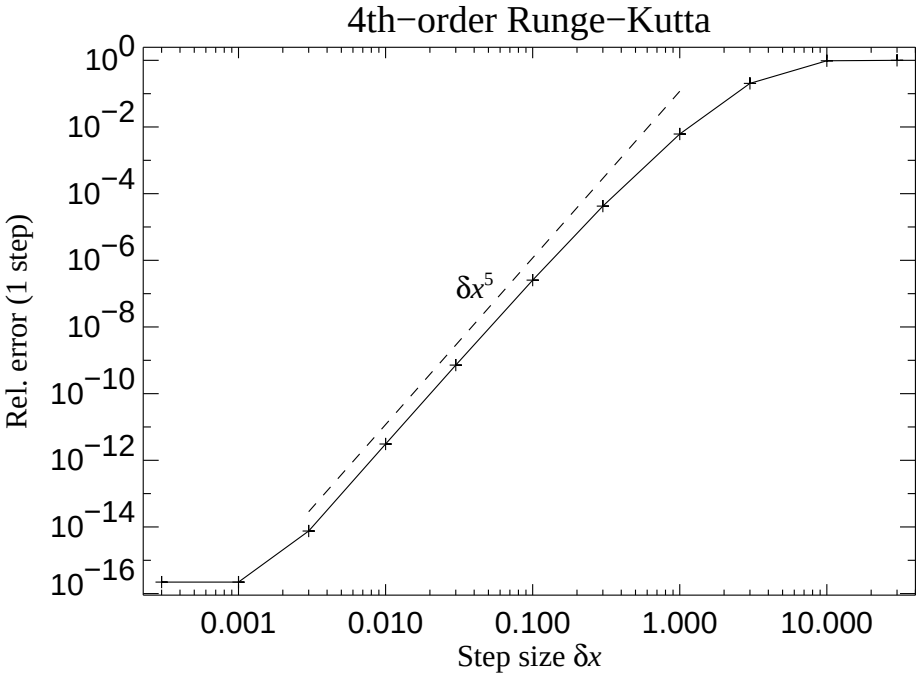

Figure 1.2: Relative error of one Runge–Kutta step applied to Eq. (1.7) [but with $t_0 = 1$] as a function of the time step δt. The dashed line shows a dependency $\sim \delta t^5$ for comparison.

t	y	
t_0	y_0	$k_1 = \delta t F(y_0, t_0)$
$t_1 = t_0 + \frac{1}{2}\delta t$	$y_1 = y_0 + \frac{1}{2}k_1$	$k_2 = \delta t F(y_1, t_1)$
$t_2 = t_0 + \frac{1}{2}\delta t$	$y_2 = y_0 + \frac{1}{2}k_2$	$k_3 = \delta t F(y_2, t_2)$
$t_3 = t_0 + \delta t$	$y_3 = y_0 + k_3$	$k_4 = \delta t F(y_3, t_3)$
$t = t_0 + \delta t$	$y = y_0 + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 + O(\delta t^5)$	

Figure 1.2 shows the relative error

$$\delta y_{\mathrm{rel}} \equiv \left| \frac{y_{\mathrm{RK}} - y_{\mathrm{exact}}}{y_{\mathrm{exact}}} \right|$$

for one time step of the classical 4th-order Runge–Kutta scheme applied to our reference problem (1.7), but with the initial condition $t_0 = 1$, $y_0 = e^{t_0} - 1 - t_0$.

Another fourth-order scheme. The classical scheme given above is by no means the only scheme involving four evaluations of the right-hand side. Here's another one; none of the schemes of a given order is *a priory* better than another, and schemes may be optimised for special requirements.

t	y	
t_0	y_0	$k_1 = \delta t F(y_0, t_0)$
$t_1 = t_0 + \frac{1}{3}\delta t$	$y_1 = y_0 + \frac{1}{3}k_1$	$k_2 = \delta t F(y_1, t_1)$
$t_2 = t_0 + \frac{2}{3}\delta t$	$y_2 = y_0 - \frac{1}{3}k_1 + k_2$	$k_3 = \delta t F(y_2, t_2)$
$t_3 = t_0 + \delta t$	$y_3 = y_0 + k_1 - k_2 + k_3$	$k_4 = \delta t F(y_3, t_3)$
$t = t_0 + \delta t$	$y = y_0 + \frac{1}{8}k_1 + \frac{3}{8}k_2 + \frac{3}{8}k_3 + \frac{1}{8}k_4 + O(\delta t^5)$	

1.3.1 Embedded Runge–Kutta schemes

For applications which require high accuracy (like some of the N-body problems discussed in Chapter 2), tight control over the time step is necessary to ensure the integration remains sufficiently accurate even when dramatic things (like close encounters of two bodies) happen. This can be achieved by so-called *embedded Runge–Kutta schemes*. For details, see Press et al. (1992, 1996) or Pozrikidis (1998)[1].

[1]But beware of a typo in one of the coefficients

Chapter 2

The 3-Body Problem

2.1 The N-body problem

Consider the motion of N point masses m_i subject to no other forces than their mutual gravity, described by Newton's gravity law. The equation of motion for the i-th particle is

$$\frac{d^2\mathbf{x}_i}{dt^2} = -G \sum_{\substack{j=1 \\ j\neq i}}^{N}{}' m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3} \ . \tag{2.1}$$

For $N = 2$, the solution of this system of equations is given by Keplerian motion of the two masses: Relative to their centre of mass, they move on conic sections and their orbits are either bound or unbound, depending on the sign of total energy.

For $N \geq 3$, only quite special solutions are known. While substantial interest in the long-term fate of the Solar system has produced elaborate perturbation methods and Lagrange's famous quote "Sire, je n'avais pas besoin de cette hypothèse," even the stability of this billion-year old N-body system has been questioned. Today, most of the investigations in N-body celestial mechanics are carried out by numerical integration of the system (2.1), be it directly, be it the numerical integration of equations for orbital elements, which are based on perturbation analysis.

2.1.1 Degrees of freedom

Historically, the case $N = 3$ has attracted much attention, culminating in a prize offered by the King of Sweden and Norway in 1887 which was won by Poincaré for showing that the problem cannot be solved in closed form — at least not in a manner analogous to the solution of the two-body problem. In fact, series solutions have been given for the three-body problem, but none of them is of any use for practical purposes.

So what makes three bodies so much more difficult to describe than two? The N-body problem possesses $N \times 3 \times 2$ degrees of freedom, resulting from motion in

three dimensions, described by second-order differential equations. For $N = 2$, this implies 12 degrees of freedom, while for $N = 3$ there are already 18.

On the other hand, this multi-dimensional phase space is not fully accessible to the system, its trajectory being restricted to hypersurfaces defined by the values of certain constants of motion. The obvious constants of motion (discussed in more detail in § 2.1.2 below) are

i) energy (1),

ii) momentum (3) and (initial) position of the common centre of mass (3),

iii) angular momentum (3),

listed here together with the number of conditions they impose. Together they neutralise 10 degrees of freedom and for the two-body problem there remain just two. The resulting second-order problem can be then be solved analytically; alternatively, one more degree of freedom can be eliminated by the *Lenz–Runge vector* as eleventh constant of motion.[1]

For the three-body problem, on the other hand, the ten invariants still leave eight degrees of freedom. And, as was shown by Bruns and Poincaré, there are no other (independent) invariants which are algebraic or even analytic in $\mathbf{x}_i$ and $\dot{\mathbf{x}}_i$ (apart from the *Jacobi invariant* in the case of the restricted three-body problem, see below). Thus, there is no hope that the three-body problem can be solved in a manner analogous to the two-body problem.

2.1.2 Conserved quantities

In this section, we derive the ten conserved quantities mentioned above for the N-body problem. These invariants are connected to fundamental symmetries of the laws of mechanics and are conserved for arbitrary central (potential) forces. Nevertheless it is instructive to derive them for the special case of the $1/r$ potential.

Energy
The potential energy of the N-body system is

$$E_{\mathrm{pot}} = -\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N}{}' \frac{G\,m_i m_j}{|\mathbf{x}_i - \mathbf{x}_j|} \;, \tag{2.3}$$

[1] For the two-body problem, the Lenz–Runge vector is given by

$$\mathbf{R} = \mathbf{L} \times (\dot{\mathbf{x}}_2 - \dot{\mathbf{x}}_1) + G m_1 m_2 \frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_2 - \mathbf{x}_1|} \;, \tag{2.2}$$

where $\mathbf{L} = m_1 \mathbf{x}_1 \times \dot{\mathbf{x}}_1 + m_2 \mathbf{x}_2 \times \dot{\mathbf{x}}_2$ denotes total angular momentum. Although all three components of $\mathbf{R}$ are conserved, only its azimuth in the plane perpendicular to $\mathbf{L}$ yields an additional constraint (the modulus $|\mathbf{R}|$, e. g., can be expressed in terms of the other constants of motion). Unlike the other ten invariants, the Lenz–Runge vector is specific to the $1/r$ potential.

so the rate of change of $E_{\rm pot}$ is, by virtue of the product rule,

$$\frac{d}{dt}E_{\rm pot} = -\frac{1}{2}{\sum_{i,j}}' G\, m_i m_j \left(\frac{d\mathbf{x}_i}{dt}\cdot\nabla_i \frac{1}{|\mathbf{x}_i-\mathbf{x}_j|} + \frac{d\mathbf{x}_j}{dt}\cdot\nabla_j \frac{1}{|\mathbf{x}_i-\mathbf{x}_j|} \right) . \tag{2.4}$$

Interchanging $i \longleftrightarrow j$, we see that the two terms in the bracket are identical; moreover, we know from potential theory that

$$\nabla\frac{1}{|\mathbf{x}|} = -\frac{\mathbf{x}}{|\mathbf{x}|^3} \quad \text{and thus} \quad \nabla_i \frac{1}{|\mathbf{x}_i-\mathbf{x}_j|} = -\frac{\mathbf{x}_i-\mathbf{x}_j}{|\mathbf{x}_i-\mathbf{x}_j|^3} , \tag{2.5}$$

so we can write

$$\frac{d}{dt}E_{\rm pot} = {\sum_{i,j}}' G\, m_i m_j \mathbf{v}_i \cdot \frac{\mathbf{x}_i-\mathbf{x}_j}{|\mathbf{x}_i-\mathbf{x}_j|^3} . \tag{2.6}$$

From the equation of motion (2.1), we can obtain a similar expression by multiplying that equation by $m_i\mathbf{v}_i$, which yields

$$m_i\mathbf{v}_i\cdot\dot{\mathbf{v}}_i = \frac{d}{dt}\left(\frac{m_i}{2}\mathbf{v}_i^2\right) = -G{\sum_{j}}' m_i m_j \mathbf{v}_i \cdot \frac{\mathbf{x}_i-\mathbf{x}_j}{|\mathbf{x}_i-\mathbf{x}_j|^3} . \tag{2.7}$$

Summing this up over all indices i, we find that

$$\frac{dE_{\rm kin}}{dt} = \frac{d}{dt}\sum_i \frac{m_i}{2}\mathbf{v}_i^2 = -\frac{d}{dt}E_{\rm pot} , \tag{2.8}$$

thus

$$E_{\rm tot} = E_{\rm kin} + E_{\rm pot} = \text{const} \tag{2.9}$$

Motion of the centre of mass; momentum
The position of the centre of mass is

$$\mathbf{X} \equiv \frac{1}{M}\sum_{i=1}^{N} m_i\mathbf{x}_i , \qquad \text{where} \qquad M \equiv \sum_{i=1}^{N} m_i \tag{2.10}$$

is the total mass of the system. Taking the second time derivative, we find that

$$M\frac{d^2\mathbf{X}}{dt^2} = \sum_i m_i\frac{d\mathbf{v}_i}{dt} = -G{\sum_{\substack{i,j=1\\ i\neq j}}^{N}}' m_i m_j \frac{\mathbf{x}_i-\mathbf{x}_j}{|\mathbf{x}_i-\mathbf{x}_j|^3} = \mathbf{0} , \tag{2.11}$$

because the expression subject to summation changes sign when i and j are interchanged.

Integrating this twice, we get

$$\mathbf{X}(t) = \mathbf{X}_0(t) + \frac{\mathbf{P}}{M}t , \tag{2.12}$$

where

$$\mathbf{X}_0 = \mathbf{const} \tag{2.13}$$

is the initial position of the centre of mass and

$$\mathbf{P} = M\frac{d\mathbf{X}}{dt} = \sum_i m_i \mathbf{v}_i = \mathbf{const} \tag{2.14}$$

is the total momentum of the system.

Angular momentum
The total angular momentum of the system is

$$\mathbf{L} = \sum_{i=1}^{N} m_i \mathbf{x}_i \times \mathbf{v}_i \tag{2.15}$$

and has the time derivative

$$\frac{d\mathbf{L}}{dt} = \sum_i m_i \left(\dot{\mathbf{x}}_i \times \mathbf{v}_i + \mathbf{x}_i \times \dot{\mathbf{v}}_i\right) = \sum_i m_i \mathbf{x}_i \times \dot{\mathbf{v}}_i \ , \tag{2.16}$$

because $\dot{\mathbf{x}}_i \times \mathbf{v}_i = \mathbf{v}_i \times \mathbf{v}_i$ vanishes.

From the equation of motion (2.1), we get

$$m_i \mathbf{x}_i \times \dot{\mathbf{v}}_i = -G \sum_{\substack{j=1 \\ j\neq i}}^{N}{}' m_i m_j \frac{\mathbf{x}_i \times (\mathbf{x}_i - \mathbf{x}_j)}{|\mathbf{x}_i - \mathbf{x}_j|^3} = G \sum_j{}' m_i m_j \frac{\mathbf{x}_i \times \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3} \ . \tag{2.17}$$

Summation over i yields

$$\sum_{i=1}^{N} m_i \mathbf{x}_i \times \dot{\mathbf{v}}_i = G \sum_{ij}{}' m_i m_j \frac{\mathbf{x}_i \times \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^3} = \mathbf{0} \ , \tag{2.18}$$

because once again the expression under the summation sign is antisymmetric in (i, j). Thus

$$\mathbf{L} = \mathbf{const} \ . \tag{2.19}$$

2.2 Special solutions

In spite of what has been said so far, all hope must not be abandoned for the three-body problem. While the general problem cannot be solved analytically, certain special cases *can*, as has been shown by L. Euler and J. L. Lagrange who were looking for form-invariant solutions, i. e. constellations for which the relative distances $R_{ij} \equiv |\mathbf{x}_i - \mathbf{x}_j|$ have common time dependence,

$$R_{ij}(t) = R_{ij}(0) f(t) \tag{2.20}$$

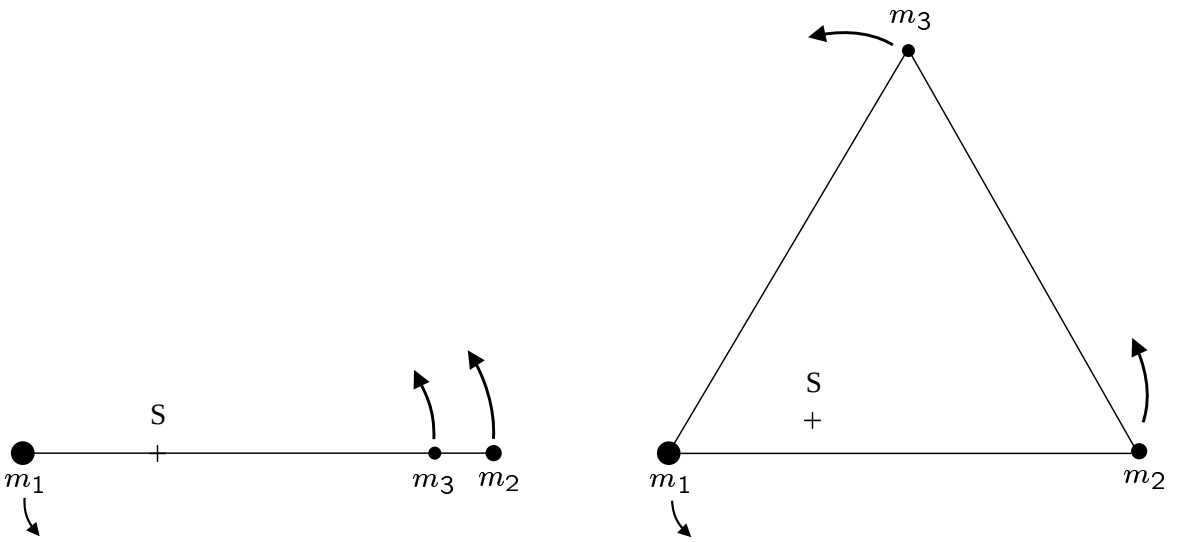

Figure 2.1: Sketch of the special solutions of Euler (collinear solution, left panel) and Lagrange (triangular solution, right panel). The three masses are $m_1 < m_2 < m_3$; the common centre of mass is indicated by the letter S.

and thus all angles remain the same.

They found two such solutions, the three masses being arranged either collinear, or as an equilateral triangle; see Fig. 2.1 for an illustration. These solutions exist for arbitrary mass ratios $m_1 : m_2 : m_3$ and for non-circular orbits. For the sake of simplicity, we will however only discuss them for the *restricted three-body problem*, where one of the masses is negligible compared to the other two, and we will focus on circular orbits. We will call the body with largest mass (m_1) the *Sun*, and the second heavy body *Jupiter* — although we will take the freedom to vary the mass ratio m_2/m_1.

Hill found another class of solutions where two of the bodies form a close binary system and interact with the distant third as one body. This configuration is always stable and is used to describe the system Earth–Moon–Sun. More generally, all stable N-body configurations with similar masses found in nature seem to have this hierarchical structure of tight two-body systems interacting with other two-body systems like single masses. However, we will now focus on Euler's and Lagrange's solutions.

2.2.1 The restricted, circular three-body problem

Under the assumptions that $m_3 \ll m_2 \leq m_1$ and that m_1 and m_2 move on circular orbits in the x-y plane, there are 5 points (*libration points*, i. e. equilibrium points) where the massless body m_3 is in force equilibrium (see Fig. 2.2). Three of these (labelled L1, L2 and L3 in Fig.2.2) represent Euler's collinear solution. The existence of these equilibrium points is pretty obvious, but they are always unstable. The other two (L4 and L5) are far less obvious, but they can be stable, as we will discuss below. They correspond to the three bodies moving as an equi-

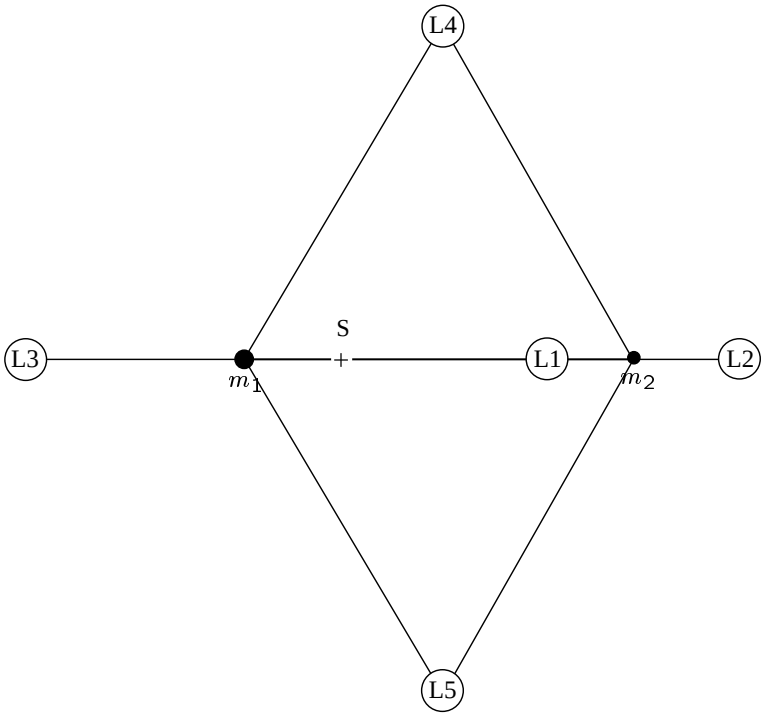

Figure 2.2: The five Lagrangian libration points L1–L5 for the restricted three-body problem.

lateral triangle. If m_2/m_1 is relatively small, the massless particle will move on Jupiter's orbit, but is 60° ahead (L4) or behind (L5) Jupiter itself. Celestial bodies in these positions have indeed been found; they are asteroids and are referred to as *Trojans*. In fact, there is a finer distinction between *Greeks* (the ones ahead of Jupiter) and the proper *Trojans* (the ones lagging behind), but we will ignore this, even more so since this naming scheme has not been applied strictly to the individual asteroids.

2.2.2 Effective potential and Jacobi's integral

We consider the circular, restricted 3-body problem. In this special case, the two heavy masses are in rigid rotation around the centre of mass with constant angular velocity

$$\omega = \sqrt{\frac{G(m_1+m_2)}{a^3}} \,. \tag{2.21}$$

It is natural to analyse the motion of m_3 in a coordinate system corotating with m_1 and m_2, with the centre of mass as origin and the z axis parallel to the angular momentum vector. In this reference frame inertial forces appear in the equation of motion for m_3, which reads

$$\ddot{\mathbf{x}} = -Gm_1\frac{\mathbf{x}_{13}}{r_{13}^3} - Gm_2\frac{\mathbf{x}_{23}}{r_{23}^3} \underbrace{+\,\omega^2 s\mathbf{e}_s}_{\mathbf{F}_{\text{cent}}} \underbrace{-\,2\vec{\omega}\times\dot{\mathbf{x}}}_{\mathbf{F}_{\text{Cor}}} \,, \tag{2.22}$$

where $s := \sqrt{x^2+y^2}$ is the cylindrical radius and $\mathbf{e}_s$ the corresponding unit vector, $\mathbf{x}_{i3} \equiv \mathbf{x}_3 - \mathbf{x}_i$, $r_{i3} \equiv |\mathbf{x}_{i3}|$, and $\vec{\omega} = (0,0,\omega)$ is the vector of angular velocity. The

last two terms in Eq. (2.22) are the centrifugal and Coriolis force, respectively. Equation (2.22) can be written as

$$\ddot{\mathbf{x}} + 2\vec{\omega} \times \dot{\mathbf{x}} = -\nabla U \ , \tag{2.23}$$

where

$$U(x, y, z) \equiv -\frac{Gm_1}{r_{13}} - \frac{Gm_2}{r_{23}} - \frac{1}{2}\omega^2 s^2 \tag{2.24}$$

is the *effective potential* of the test mass m_3.[2]

Multiplying Eq. (2.23) by $\mathbf{v}$, we obtain

$$\frac{d}{dt}\frac{\mathbf{v}^2}{2} = -\frac{d}{dt}U(x, y, z) \ , \tag{2.28}$$

which implies that

$$J(\mathbf{x}, \dot{\mathbf{x}}) \equiv \frac{\dot{\mathbf{x}}^2}{2} + U(\mathbf{x}) = \text{const.} \tag{2.29}$$

The quantity J is called *Jacobi integral* and it represents the eleventh integral of motion for the restricted, circular three-body problem. As was shown by Poincaré, no further algebraic integral of motion exists for this special case of the three-body problem.

From this point on, we will consider the planar (restricted, circular) problem only, i. e. the case where $z = 0$. This is a very natural special case, as the z-dependence of the effective potential $U(\mathbf{x})$ is such that there is always a restoring force towards the x-y plane.

For a given trajectory, J is constant. Hence, all points where $|\mathbf{v}| = 0$ (if they exist) are located on lines $U = J = \text{const}$. Figure 2.4 shows lines of constant potential U in the plane $z = 0$. Depending on the value of J, the body m_3 can only access the region where $U(\mathbf{x}) \leq J$; it cannot cross the line $U = J$, because otherwise $\mathbf{v}$ would become imaginary. This implies that if J is sufficiently small, the test mass will be confined to the vicinity of either m_1 or m_2 forever, which is in accordance with Hill's result and is illustrated by the stability of planet–satellite systems.

[2] In Cartesian coordinates this would be

$$\ddot{x} - 2\omega\dot{y} = -\frac{\partial U}{\partial x} \ , \tag{2.25}$$

$$\ddot{y} + 2\omega\dot{x} = -\frac{\partial U}{\partial y} \ , \tag{2.26}$$

$$\ddot{z} = -\frac{\partial U}{\partial z} \ , \tag{2.27}$$

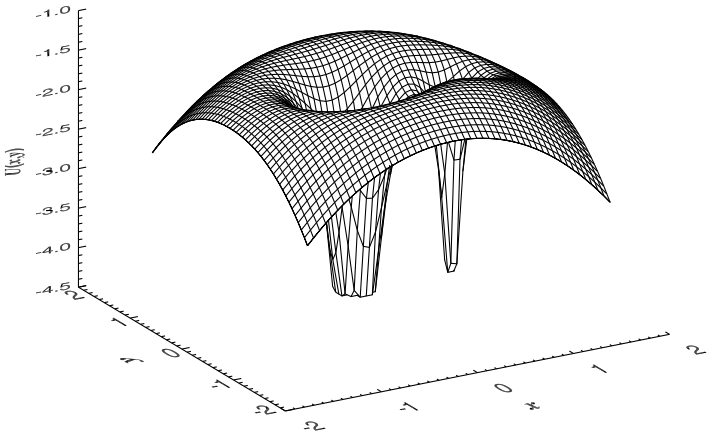

Figure 2.3: Surface plot of the effective potential $U(x, y)$ for the planar, restricted, circular three-body problem; masses are $m_1 = 0.8, m_2 = 0.2, m_3 = 0$.

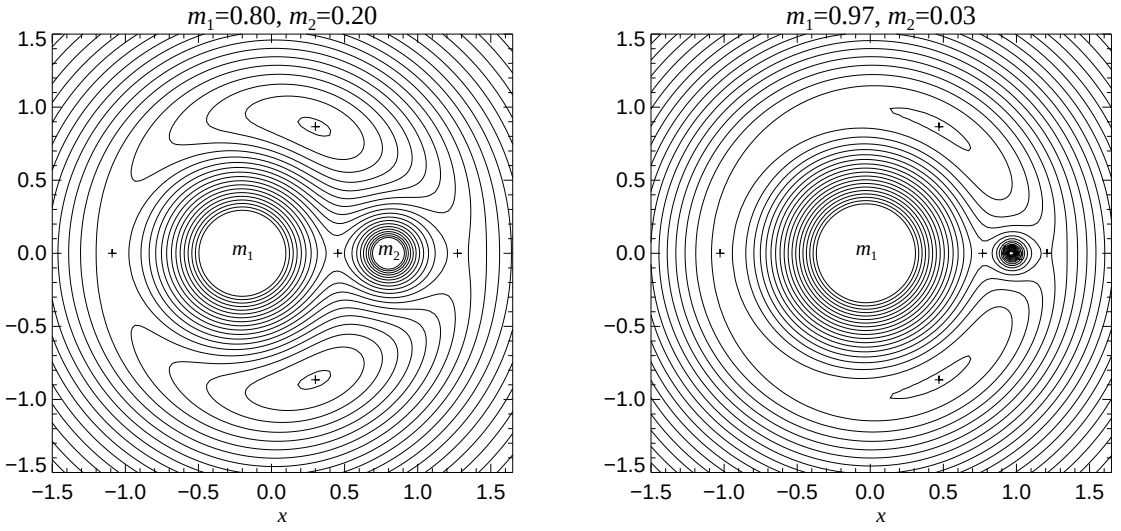

Figure 2.4: Contour lines of the effective potential U(x,y) for the planar, restricted, circular three-body problem; Left: masses $m_1 = 0.8, m_2 = 0.2$; right: $m_1 = 0.97, m_2 = 0.03$; in both cases $m_3 = 0$. The five libration points are indicated by crosses; the centre of mass is at $(0, 0)$.

2.2.3 Stability of libration points

Figure 2.3 shows that none of the libration points corresponds to a minimum of the effective potential U: L1–L3 are saddle points, while L4 and L5 are maxima of U. Thus, none of these points seems to allow a stable equilibrium. It is easy to show generally that the effective potential $U(\mathbf{x})$ of a rigidly rotating mass distribution cannot have a minimum in a point where density ϱ vanishes, i. e. outside the gravitating bodies. To see this, we recall that the gravity potential $\Phi(\mathbf{x})$ satisfies Poisson's equation

$$\Delta\,\Phi = -4\pi G\varrho\ . \tag{2.30}$$

In a reference frame rotating at angular momentum ω, the effective potential is

$$U(\mathbf{x}) = \Phi(\mathbf{x}) - \frac{1}{2}\omega^2 s^2\ ; \tag{2.31}$$

the Laplacian of U is

$$\Delta\,U = \Delta\,\Phi - \frac{\omega^2}{2}\,\Delta\,s^2 = \Delta\,\Phi - 2\omega^2\ . \tag{2.32}$$

Now if $\varrho = 0$, the Laplacian $(\partial_{xx}+\partial_{yy}+\partial_{zz})U = -2\omega^2$ is negative and thus at least one of the second derivatives $\partial_{xx}U$, $\partial_{yy}U$, $\partial_{zz}U$ must be negative, which implies instability in the corresponding direction.

This does however not mean that none of the libration points can be stable. The Coriolis force — which did not enter this analysis at all — can play an important role in stabilising a configuration by preventing m_3 from "rolling down the slope of the potential". A conclusive stability investigation will be based on linear stability analysis of Eq. (2.23) by making the ansatz

$$(x, y, z)(t) = e^{\gamma t}(\hat{x}, \hat{y}, \hat{z})\ , \tag{2.33}$$

linearising the potential U around the equilibrium point and solving the resulting linear eigenvalue problem for the growth rate γ. If there exists an eigenvalue γ with positive real part, then the corresponding point is unstable, otherwise it is (linearly) stable.

We will not carry out this analysis here, but just report that L1, L2 and L3 turn out to be always unstable, while L4 and L5 are stable, provided that

$$m_2/(m_1{+}m_2) < \frac{1-\sqrt{23/27}}{2} \approx 0.0385\ , \tag{2.34}$$

or $m_2/m_1 < (25-\sqrt{621})/2 \approx 0.0401$, and unstable if m_2/m_1 is larger. The mass ratio of Jupiter and the Sun is $m_{♃}/m_{☉} \approx 1.9{\times}10^{27}\,\mathrm{kg}/2.0{\times}10^{30}\,\mathrm{kg} \approx 0.00095$; correspondingly, the orbits of the Trojans and Greeks are stable, at least as far as our assumptions are justified for this system. In reality, Jupiter's orbit is not circular and Saturn and other planets have an influence as well, so our conclusion must be taken *cum grano salis*.

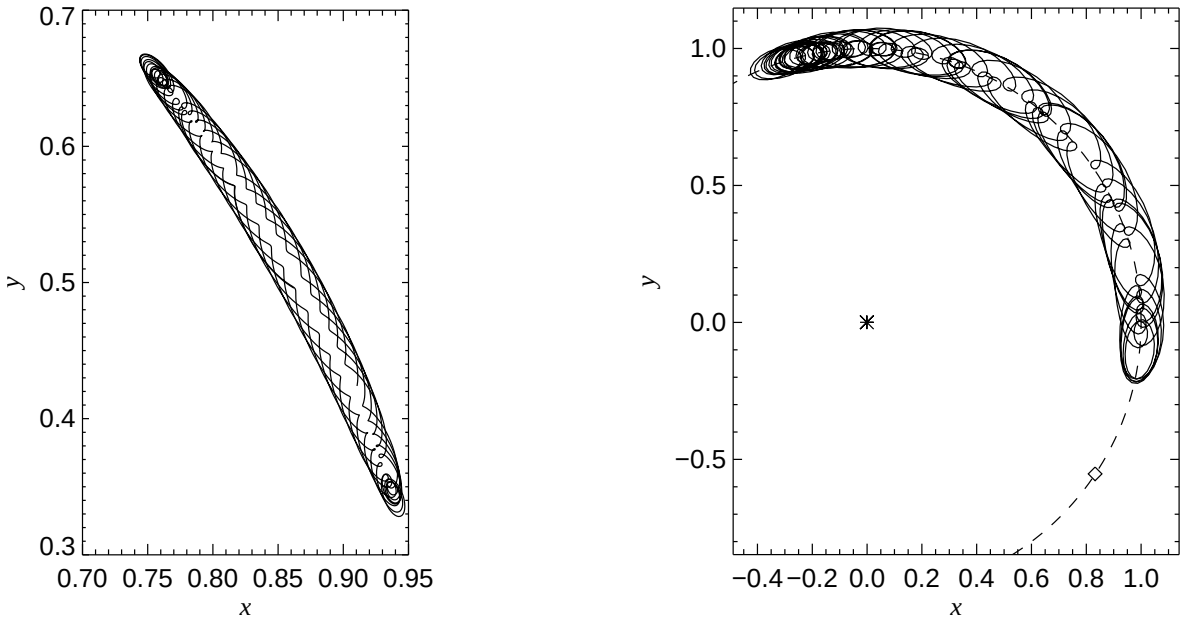

Figure 2.5: Two Trojan trajectories for $m_2/m_1 = 0.001$ in the corotating reference frame. Left: moderate perturbation relative to the equilibrium position. Right: strong perturbation. The asterisk ($*$) marks the Sun's position, the diamond ($\diamond$) marks Jupiter; Jupiter's orbit is shown as a dashed line.

2.2.4 Trajectories of Trojans

The Trojans move on complicated trajectories which (for small amplitude of the initial perturbation) are bounded by a kidney-shaped envelope. Figure 2.5 shows two sample trajectories for a mass ratio m_2/m_1 similar to that of Jupiter and the Sun. The second figure shows a trajectory with a stronger perturbation relative to the equilibrium point. Note how the trajectory fills a considerable part of Jupiter's orbit.

2.2.5 Chaos in the restricted three-body problem

If $m_1 = m_2$, the two heavy bodies will always have the same distance from the centre of mass; this special case is referred to as *Copenhagen problem*. In this case, if m_3 initially moves along the z-axis (which passes through the centre of mass) and its velocity has only a z-component, then it will remain on this axis forever, because the horizontal forces cancel for symmetry reasons. The resulting one-dimensional problem is still far from trivial and was found to be chaotic for certain eccentricities and initial conditions.

2.2.6 Recent results

New solutions to the three-body or N-body problem are still being found, see `http://www.ams.org/notices/200105/fea-montgomery.pdf`, `http://www.maia.ub.es/dsg/3body.html`, where 'Figure-Eight' solutions and `http://www.maia.ub.es/dsg/nbody.html`, where more complex "choreographies" are dis-

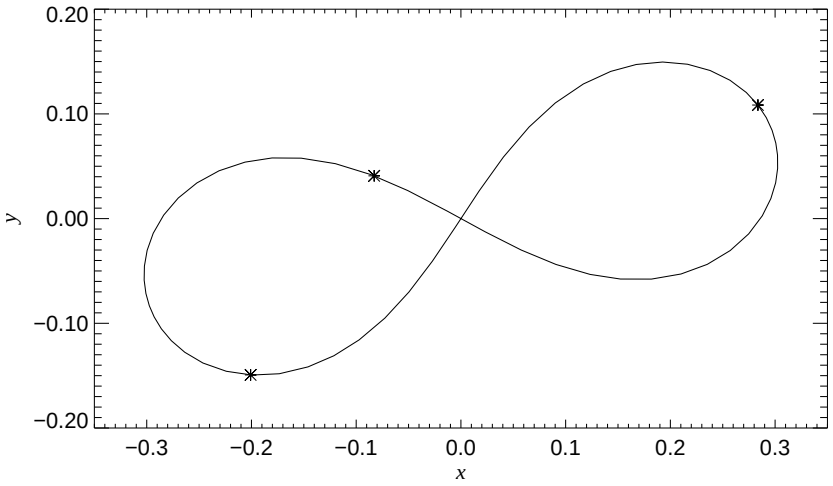

Figure 2.6: "Figure-eight" solution of the three-body problem. Parameters are $m_1 = m_2 = m_3$; the solutions has vanishing total angular momentum.

cussed. Figure 2.6 shows the Figure-eight solution (see Simó; 2002, for the corresponding initial conditions). Figures 2.6 to 2.10 show some other "choreographic" solutions taken from C. Simó's website (`http://www.maia.ub.es/dsg/nbody.html`).

These recent solutions (at least the Figure-eight one) are stable in a weak sense ("KAM-stable"). This means that most solutions with initial conditions close to the orbit stay close to it forever. The density of those few that don't tends to zero as one approaches the figure-eight solution. And the unstable trajectories diverge only very slowly from the figure eight.

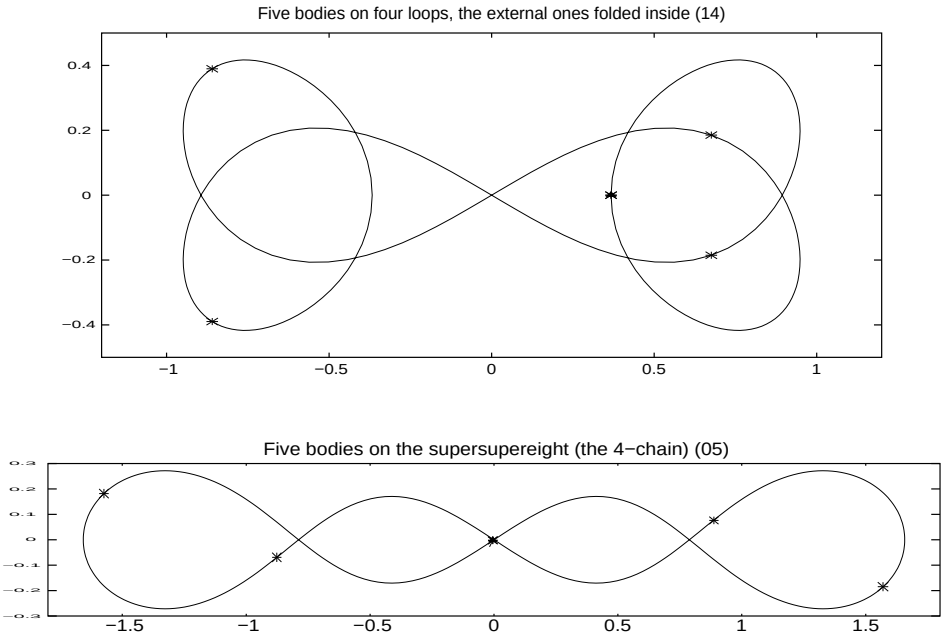

Figure 2.7: "Choreographic" solutions of the 5-body problem. Top: the '5-body 4-chain'; bottom: the '5-body 4-loop'.

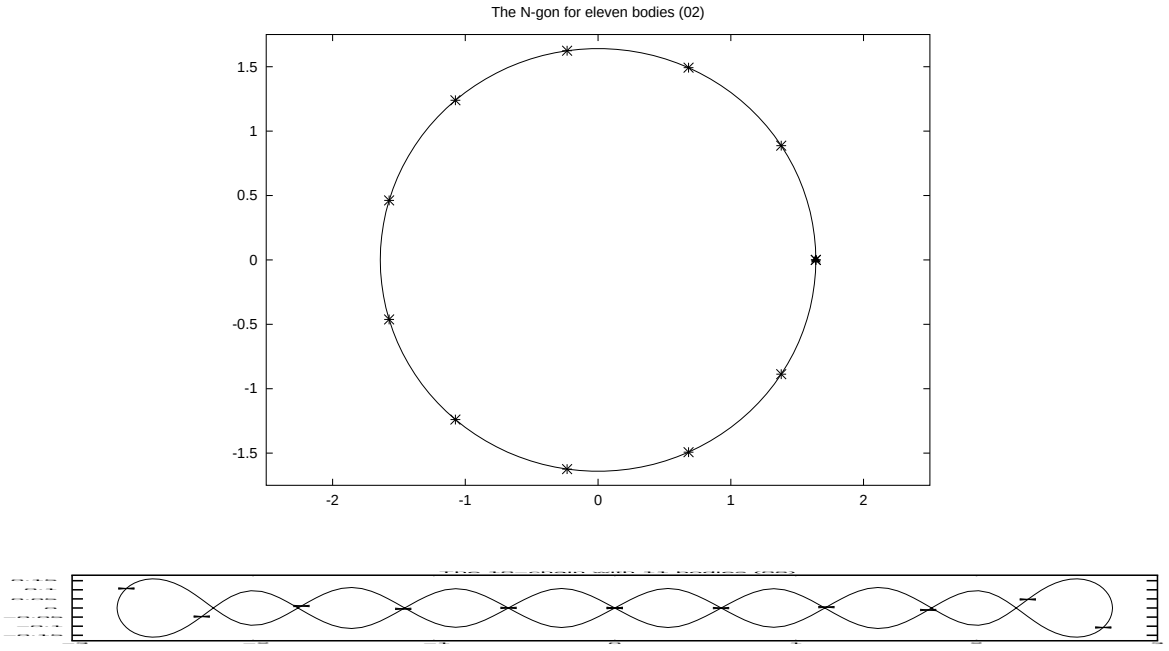

Figure 2.8: "Choreographic" solutions of the 11-body problem. Top: N-gon solution (a generalisation of Lagrange's triangular solution). Bottom: the '11-body 10-chain'; Note that this figure's aspect ratio is not correct.

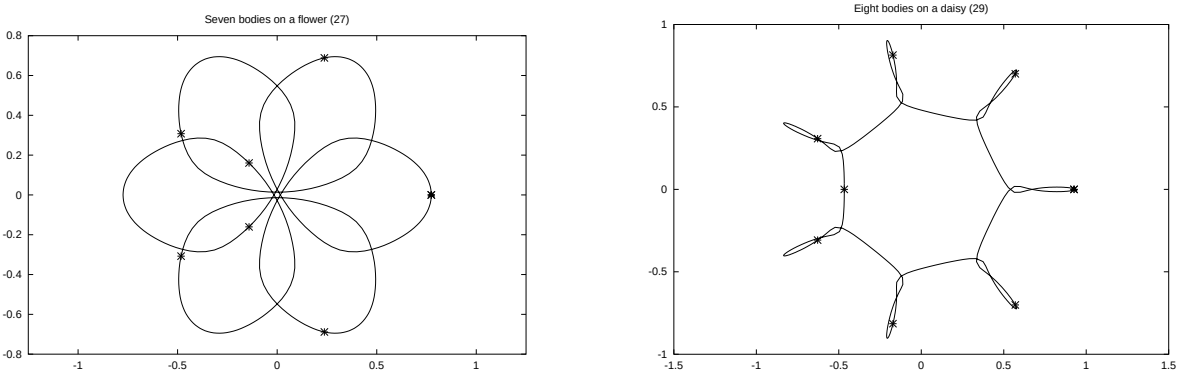

Figure 2.9: “Floral” solutions of the 7- and 8-body problem. Left: ‘7-body sunflower solution’; right: ‘8-body daisy solution’.

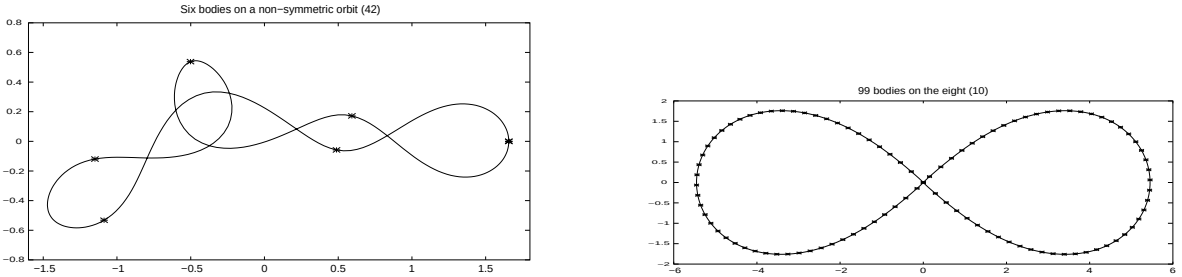

Figure 2.10: Asymmetric solution of the 6-body problem (left). 99 bodies on a figure-eight (right)

Chapter 3

Charged Particles in the Ionosphere

The equation of motion for a particle of electric charge q and mass m in a magnetic field is

$$m\frac{d^2\mathbf{x}}{dt^2} = m\frac{d\mathbf{v}}{dt} = q\mathbf{v}\times\mathbf{B}\ , \tag{3.1}$$

where $\mathrm{v} = d\mathrm{x}/dt$ is the velocity vector, and B denotes the magnetic flux density.

3.1 Point charge in a homogeneous B-field

In a homogeneous magnetic field $\mathrm{B} = B\mathrm{e}_z$, the equation of motion (3.1) becomes

$$m\frac{dv_x}{dt} = qv_yB\ , \tag{3.2}$$

$$m\frac{dv_y}{dt} = -qv_xB\ , \tag{3.3}$$

$$m\frac{dv_z}{dt} = 0\ . \tag{3.4}$$

The z-component thus decouples from x and y. Its solution describes a uniform motion. The kinematics in the plane perpendicular to the magnetic field is conveniently solved for in terms of the complex variable

$$w \equiv v_x + iv_y\ , \tag{3.5}$$

for which we obtain the equation

$$m\frac{dw}{dt} = -iqwB\ . \tag{3.6}$$

The solution is

$$w = v_\perp e^{\mp i\omega_\mathrm{L}t}\ , \tag{3.7}$$

where

$$\omega_\mathrm{L} \equiv \frac{|q|B}{m} \tag{3.8}$$

is called the *Larmor frequency*, *cyclotron frequency*, or *gyration frequency*. The upper sign is for a particle with positive charge, the lower sign for a particle with

negative charge. This solution describes a rotation of the velocity vector around the direction of B,

$$v_x = v_\perp \cos\omega_\mathrm{L}(t-t_0) \;, \tag{3.9}$$
$$v_y = \mp v_\perp \sin\omega_\mathrm{L}(t-t_0) \;, \tag{3.10}$$

where $v_\perp$ and t_0 are determined from the initial conditions ($v_\perp = (v_x^2 + v_y^2)^{1/2}$ for all t). We set $t_0 = 0$; for a convenient choice of the origin, the trajectory of the particle is then given by

$$x = x_0 + \frac{v_\perp}{\omega_\mathrm{L}} \sin\omega_\mathrm{L} t \;, \tag{3.11}$$
$$y = y_0 \pm \frac{v_\perp}{\omega_\mathrm{L}} \cos\omega_\mathrm{L} t \;, \tag{3.12}$$
$$z = z_0 + v_{z,0} t \;. \tag{3.13}$$

This describes a helix winding around a central field line. The radius

$$r_\mathrm{L} = \frac{v_\perp}{\omega_\mathrm{L}} \tag{3.14}$$

is called *Larmor radius*. The sense of the gyration is such that the electric current that arises from the moving charged particle constitutes a source for a magnetic field that is *opposite* to the original field (Fig. 3.1). Thus, charged particles in a magnetic field behave “diamagnetically”.

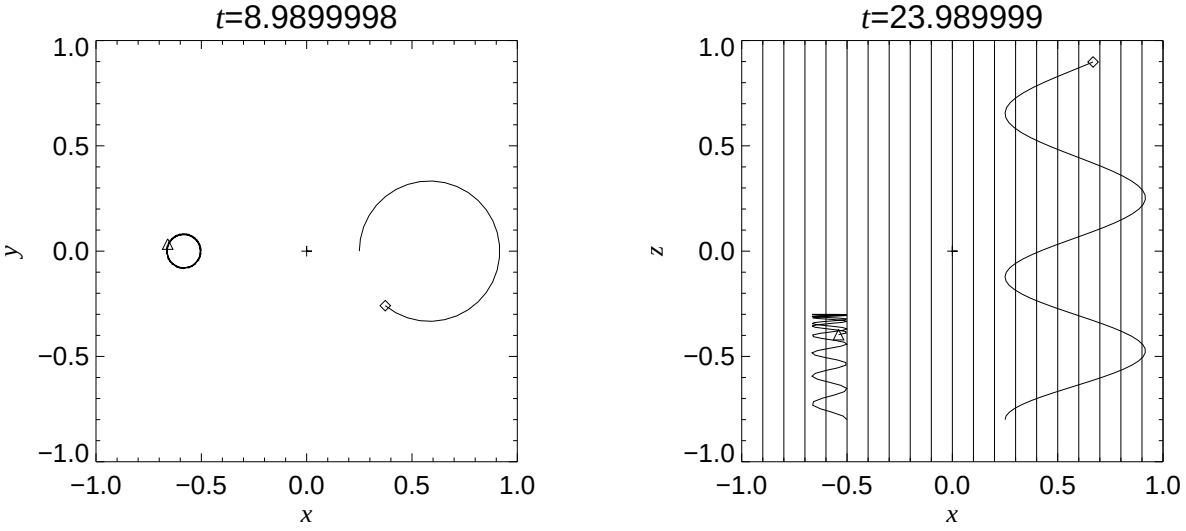

Figure 3.1: Gyration of charged particles in parallel homogeneous magnetic and electric fields. *Left*: View parallel to B (the field points toward the viewer). The left particle is negative, the right positive; the charge per mass of the left particle is 4 times larger, both particles have the same $v_\perp$. *Right*: View perpendicular to the field; the acceleration caused by **E** (in $+z$-direction) depends on the particle charge.

3.2 Homogeneous magnetic and electric fields

We take again a magnetic field in the z-direction, $\mathbf{B} = B\mathrm{e}_z$. In general, then, the electric field, $\mathbf{E}$, has components in all three directions. The equation of motion is

$$m\frac{d^2\mathbf{x}}{dt^2} = m\frac{d\mathbf{v}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \ . \tag{3.15}$$

For the z-component we have only the electric force, which yields an accelerated motion (as shown in Fig. 3.1 for the case of an electric field parallel to $\mathbf{B}$),

$$v_z = v_{z,0} + \frac{q}{m}E_z t \ , \tag{3.16}$$

and

$$z = z_0 + v_{z,0}t + \frac{q}{2m}E_z t^2 \ . \tag{3.17}$$

The x- and y-components are again combined into a complex equation; using $E_c = E_x + iE_y$ we obtain

$$m\frac{dw}{dt} = q(E_c - iwB) \ . \tag{3.18}$$

This equation has the exact solution

$$w = v_\perp e^{\mp i\omega_\mathrm{L} t} + \frac{E_c}{iB} \ , \tag{3.19}$$

or

$$v_x = v_\perp \cos\omega_\mathrm{L} t + \frac{E_y}{B} \ , \tag{3.20}$$

$$v_y = \mp v_\perp \sin\omega_\mathrm{L} t - \frac{E_x}{B} \ . \tag{3.21}$$

This solution consists of the gyration of the particle around the lines of force, as before, and – additionally – of a drift

$$\mathbf{v}_\mathrm{D} = (\frac{E_y}{B}, -\frac{E_x}{B}, 0) = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \ . \tag{3.22}$$

This drift is called the $\mathbf{E} \times \mathbf{B}$ drift. It is in the *same direction* for positive and negative particles. Notice that this time $v_\perp$ is again a constant of integration, but is *not* identical to $(v_x^2 + v_y^2)^{1/2}$. Notice also that the drift is only obtained for $B \neq 0$; if $B = 0$ then we must return to the original equation of motion which yields the well-known accelerated motion of a charged particle in an electric field. In the Earth's ionosphere, the particles gyrate in the magnetic field, and the $\mathbf{E} \times \mathbf{B}$ drift is superposed to this gyration.

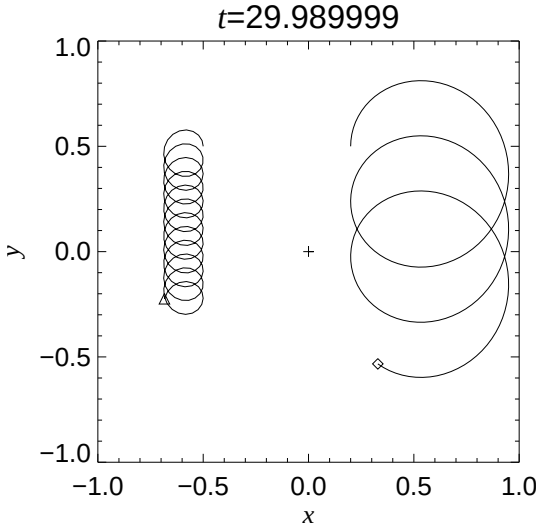

Figure 3.2: The $\mathbf{E} \times \mathbf{B}$ drift. The magnetic field is towards the viewer, the electric field in the x-direction. Thus, according to (3.22), both particles drift into the $-y$-direction, with the same speed, $-E_x/B_z$.

3.3 Inhomogeneous magnetic field

Next we consider the case where $\mathbf{E} = 0$, and $\mathbf{B}$ is in the z-direction, but with a dependence perpendicular to $\mathbf{B}$. In this case we shall obtain only an approximate solution, in contrast to the previous two sections where the equation of motion was solved exactly. We choose the coordinate system such that

$$\mathbf{B} = (\ 0,\ 0, B(y)) \tag{3.23}$$

The y-dependence of B shall be weak. Therefore we expand

$$B(y) = B_0 + (y - y_0)\frac{dB}{dy}\ , \tag{3.24}$$

where the field gradient dB/dy is a small constant in the sense $r_{\mathrm{L}}|dB/dy| \ll |B_0|$. That is, the length over which the field changes noticeably is large compared to the Larmor radius. The coefficient, $y - y_0$, of the small quantity can then be taken from the Larmor orbit (3.12) in a homogeneous field, so that

$$B = B_0 \pm r_{\mathrm{L}} \cos\omega_{\mathrm{L}} t\, \frac{dB}{dy}\ . \tag{3.25}$$

The z-component of the equation of motion yields a uniform motion, which we shall not consider further. The perpendicular components are

$$m\frac{dv_x}{dt} = qv_y B_0 \pm qv_y r_{\mathrm{L}} \cos\omega_{\mathrm{L}} t\, \frac{dB}{dy}\ , \tag{3.26}$$

$$m\frac{dv_y}{dt} = -qv_x B_0 \mp qv_x r_{\mathrm{L}} \cos\omega_{\mathrm{L}} t\, \frac{dB}{dy}\ . \tag{3.27}$$

We seek a solution of the form

$$v_x = \bar{v}_x + v_\perp \cos\omega_{\rm L} t \ , \tag{3.28}$$

$$v_y = \bar{v}_y \mp v_\perp \sin\omega_{\rm L} t \ , \tag{3.29}$$

i.e., a gyration plus a uniform mean motion. We substitute this into the two equations, and average the equations over time. There is no contribution on the left, and on the right only from the term that involves $\cos^2\omega_{\rm L} t$, with $1/2$ as its mean. Hence the result is

$$\bar{v}_x = \mp \frac{v_\perp r_{\rm L}}{2B_0}\frac{dB}{dy} \ , \tag{3.30}$$

$$\bar{v}_y = 0 \ . \tag{3.31}$$

We have obtained this result for a special choice of the coordinate system. The general form is

$$\mathbf{v}_{\rm D} = \pm \frac{v_\perp r_{\rm L}}{2B_0^2}(\mathbf{B}\times\nabla B) \ . \tag{3.32}$$

This mean uniform motion is called the $\mathbf{B}\times\nabla B$ drift. This drift is in *opposite directions* for positive (upper sign) and negative (lower sign) particles.

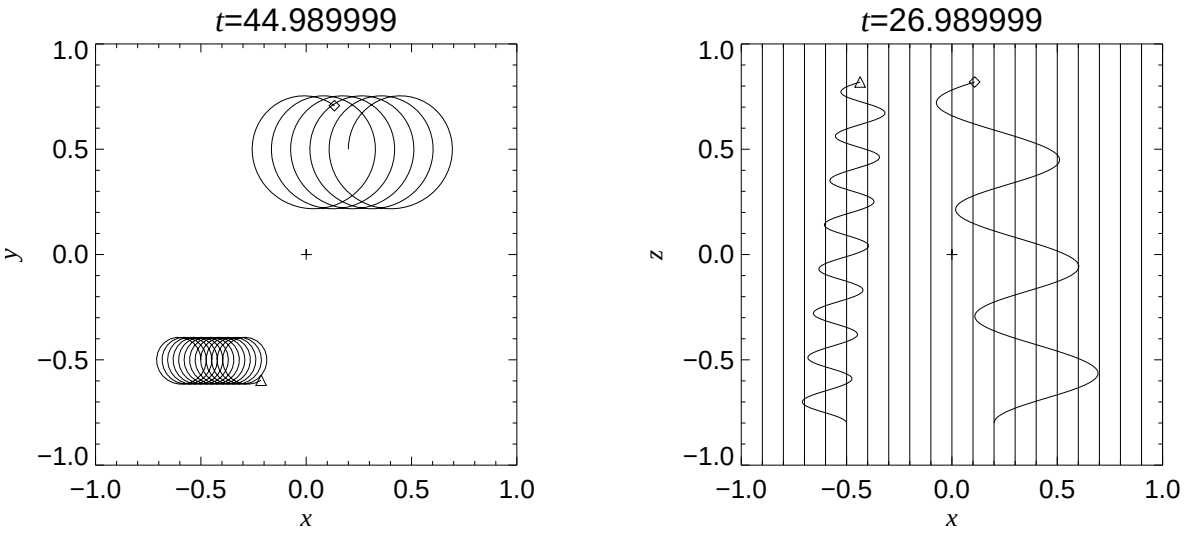

Figure 3.3: The $\mathbf{B}\times\nabla B$ drift. The magnetic field is in the positive z-direction, increasing with y. *Left*: View parallel to $\mathbf{B}$ (the field points toward the viewer). The lower particle is negative and drifts to the right, the upper is positive and drifts to the left, cf. (3.32). *Right*: View perpendicular to $\mathbf{B}$; the vertical motion is uniform and equal to the initial velocity component $v_{z,0}$.

3.4 Curvature drift

So far we have dealt with magnetic fields that have straight lines of force, although a field gradient perpendicular to the field was admitted. The Earth has a field that is dipolar in a first approximation, and the field lines of this field are curved. Particles that gyrate around such curved field lines will be subject to a

centrifugal force. This leads to a drift that can be treated in analogy to the $\mathbf{E} \times \mathbf{B}$ drift of Sect. 3.2.

Let $v_\parallel$ be the velocity component in the direction of the magnetic field, and let R_c be the radius of curvature of the field lines. The centrifugal force experienced by a particle that tries to gyrate in its helix along the field is then $mv_\parallel^2/R_c$ or, in vectorial form,

$$\mathbf{F} = \frac{mv_\parallel^2}{R_c^2}\mathbf{R_c} \ , \tag{3.33}$$

where $\mathbf{R_c}$ is the vector that points from the center of curvature to the particle. In order to obtain the curvature drift, we simply replace the electric force $q\mathbf{E}$ by $\mathbf{F}$. Hence the formula (3.22) is replaced by

$$\mathbf{v}_\mathrm{D} = \frac{mv_\parallel^2}{qB^2}\frac{\mathbf{R_c} \times \mathbf{B}}{R_c^2} \ . \tag{3.34}$$

Since in general the field lines are not circles, this result is only an approximation.

For the Earth's magnetic field the field strength decreases with increasing distance. As the field is current-free, we can estimate the field gradient from the condition $\mathrm{curl}\,\mathbf{B} = 0$: consider a system of cylindrical coordinates (s, ϕ, z) such that s is the distance from the center of curvature, ϕ an azimuth in the plane containing the field line about which the particle gyrates and the center of curvature, and z perpendicular to that plane. The z-component of the current-free condition then reads

$$\frac{1}{s}\frac{\partial}{\partial s}(sB_\phi) = 0 \ , \tag{3.35}$$

or $B_\phi \propto 1/s = 1/R_c$. On the other hand, we have, approximately, $|B| = B_\phi$; hence $|B| \propto 1/R_c$ and

$$\frac{\nabla|B|}{|B|} = -\frac{\mathbf{R_c}}{R_c^2} \ . \tag{3.36}$$

We use this to rewrite the $\mathbf{B} \times \nabla B$ drift (3.32) in the form

$$\mathbf{v}_\mathrm{D} = \pm\frac{v_\perp r_\mathrm{L}}{2B}\frac{\mathbf{R_c} \times \mathbf{B}}{R_c^2} = \frac{mv_\perp^2}{2qB^2}\frac{\mathbf{R_c} \times \mathbf{B}}{R_c^2} \ . \tag{3.37}$$

Combined with the curvature drift this yields

$$\mathbf{v}_\mathrm{D} = \frac{m}{qB^2}\left(v_\parallel^2 + \frac{1}{2}v_\perp^2\right)\frac{\mathbf{R_c} \times \mathbf{B}}{R_c^2} \ . \tag{3.38}$$

The two contributions are in the same direction. In the ionosphere, the drift is westward for the positive ions, and eastward for the negative electrons (Fig. 3.6).

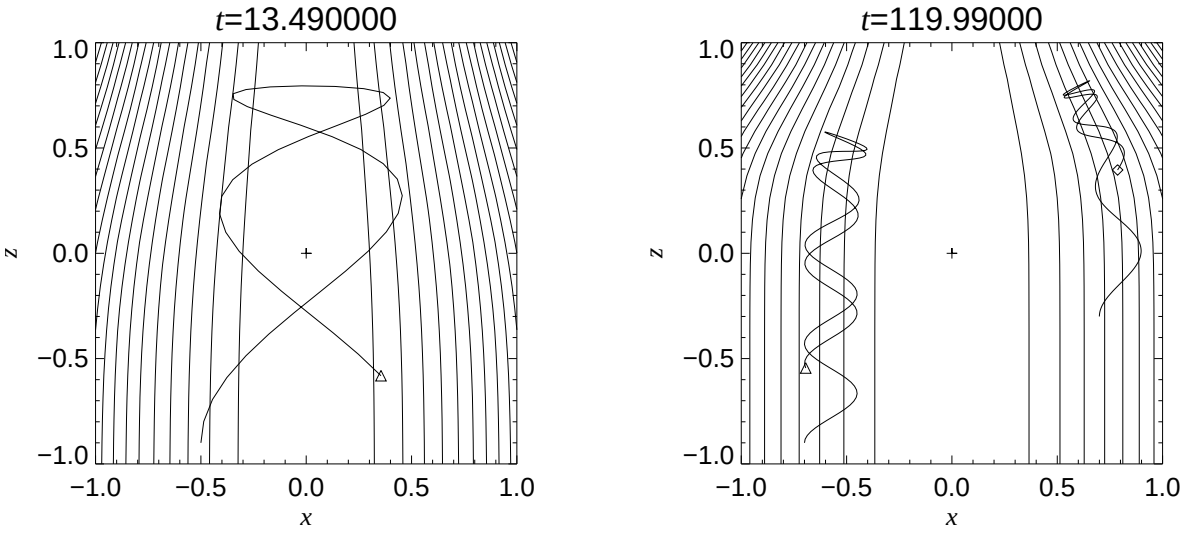

Figure 3.4: Cut through an axisymmetric magnetic mirror. The z-axis is the axis of symmetry. Although the field is nearly homogeneous at the lower boundary, the lines of force are not equidistant because between two neighbouring axisymmetric surfaces is always the *same magnetic flux*. *Left*: A particle moving around the axis of symmetry. *Right*: Two particles outside the axis of symmetry. The particles follow the field lines; the magnetic flux enclosed by their helical motion is constant.

3.5 A magnetic mirror

In this section we consider another case of a magnetic-field inhomogeneity, namely a field that varies along its own direction. Especially, the field shall be rotationally symmetric around the z-axis, and the field strength shall increase in the z-direction, so that the field lines form a kind of bottleneck. In cylindrical coordinates (s, ϕ, z) the field is

$$\mathbf{B} = (\, B_s(s,z),\, 0,\, B_z(s,z)) \; . \tag{3.39}$$

Consider a particle gyrating around the axis of symmetry. The z-component of the equation of motion is

$$m\frac{dv_z}{dt} = -qv_\phi B_s \; . \tag{3.40}$$

The equation $\mathrm{div}\,\mathbf{B} = 0$ reads

$$\frac{1}{s}\frac{\partial}{\partial s}(sB_s) + \frac{\partial B_z}{\partial z} = 0 \; , \tag{3.41}$$

which can be used to express B_s in terms of B_z,

$$B_s = -\frac{1}{s}\int_0^s s' \, \frac{\partial B_z}{\partial z}(s')ds' \; . \tag{3.42}$$

Now we restrict the attention to the approximate case where the field is predominantly in the z-direction, i.e., $B_z \approx B$, with $\partial B/\partial z \approx$ const.. Then $B_s \approx -(s/2)\partial B/\partial z$, which we substitute into the equation of motion:

$$m\frac{dv_z}{dt} = \frac{qv_\phi s}{2}\frac{\partial B}{\partial z} = \mp\frac{qv_\perp r_\mathrm{L}}{2}\frac{\partial B}{\partial z} = -\frac{mv_\perp^2}{2B}\frac{\partial B}{\partial z} \; . \tag{3.43}$$

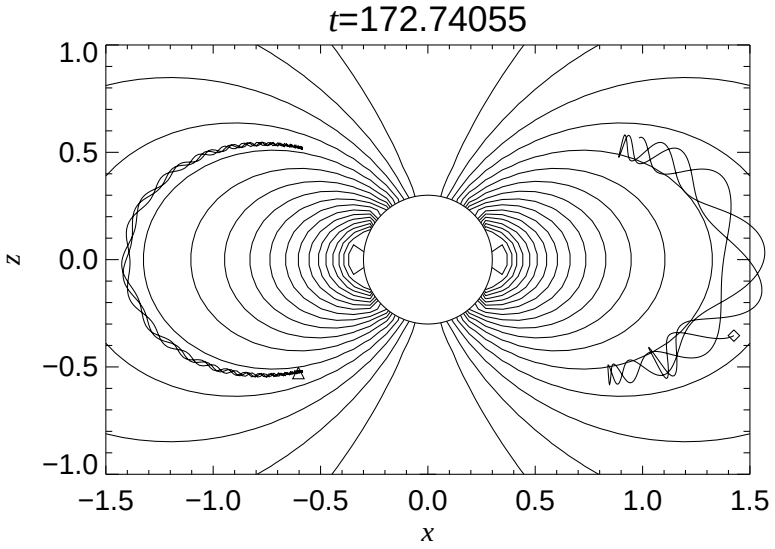

Figure 3.5: Two particles that are trapped in a dipole magnetic field. The two orbits are projected onto meridional planes $180°$ apart (left and right, respectively). Both particles are injected at the equator with the same perpendicular velocity $v_\perp$, but the particle on the left has a charge/mass value that is three times that of the right particle.

Again the upper sign is for positive particles (where $v_\phi = -v_\perp$), and the lower sign is for negative particles (where $v_\phi = v_\perp$). We have used here $s = r_\mathrm{L}$, because a gyration around the axis of symmetry was considered.

Now we define the quantity

$$\mu = \frac{m v_\perp^2}{2B} \,, \tag{3.44}$$

which is called the *magnetic moment* of the gyrating particle. We shall show presently that, as the particle gyrates gradually into the bottleneck formed by the magnetic field, μ remains approximately constant (the *adiabatic invariant*, cf. Sect. 3.6). In this case the equation of motion has an integral,

$$v_z = v_{z,0} - \frac{\mu}{m}\frac{\partial B}{\partial z}\, t \,. \tag{3.45}$$

Since $\partial B/\partial z > 0$, the velocity v_z into the bottleneck decreases, and the windings of the helix become flatter. Its is even possible that v_z changes its sign, i.e., the particle is reflected from the bottleneck. For this reason the field considered in this section is also called a *magnetic mirror*. Imagine a field with two such mirrors facing each other. The charged particles may be reflected back and forth between these two mirrors, and so are effectively enclosed in the field, which is also called a *magnetic bottle*.

The terrestrial magnetic field is approximately dipolar. Its strength increases towards both magnetic poles. Hence there are two mirrors, although not directly face to face (Fig. 3.5). Between the two mirrors the field is both inhomogeneous (in the perpendicular direction) and curved. Therefore the particles enclosed in the bottle are subject to the $\mathrm{B} \times \nabla B$ drift as well as to the curvature drift.

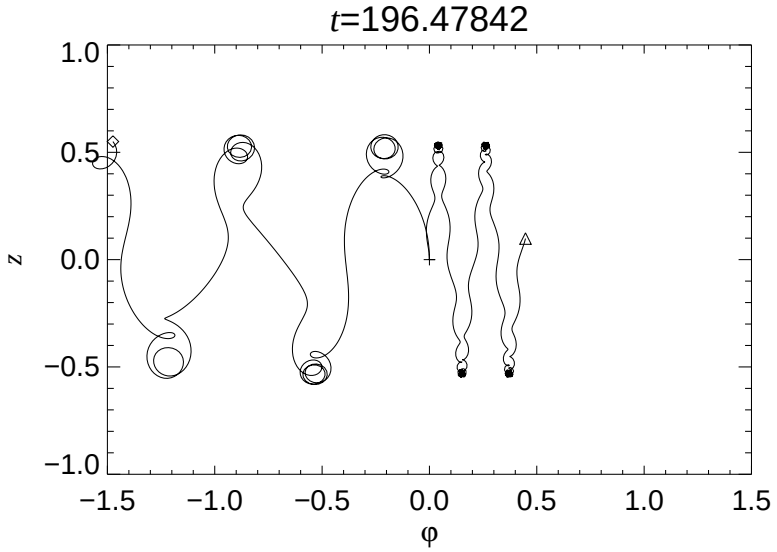

Figure 3.6: Two particles as in Fig. 3.5. The orbits are projected onto a cylindrical surface around the Earth. The particle with the positive charge drifts westwards (*left*), the negative particle drifts eastwards (*right*). Notice that magnetic north is at geographic south.

If the magnetic mirror consists of a field inhomogeneity that moves (with the motion of a cloud in interstellar space, say), then the reflected particles can be accelerated and gain energy from that motion. This is called the *Fermi mechanism*, and has been proposed as a mechanism that leads to the high energies of charged cosmic-ray particles.

3.6 Adiabatic invariants

3.6.1 The magnetic moment

The first adiabatic invariant is the magnetic moment μ that we have already used in the preceding section. In order to show that μ is constant we return to the equation of motion, which we write slightly more general by replacing v_z by $v_\parallel$:

$$m\frac{dv_\parallel}{dt} = -\mu\,\frac{\partial B}{\partial z}\ . \tag{3.46}$$

Now $v_\parallel = dz/dt$; multiplication of (3.46) by $v_\parallel$ therefore yields

$$mv_\parallel\frac{dv_\parallel}{dt} = \frac{m}{2}\frac{dv_\parallel^2}{dt} = -\mu\,\frac{dB}{dt}\ , \tag{3.47}$$

or, since $v_\parallel^2 + v_\perp^2$ is a constant (the magnetic force is always perpendicular to the motion and hence does not change the kinetic energy),

$$\frac{m}{2}\frac{dv_\perp^2}{dt} = \mu\,\frac{dB}{dt}\ . \tag{3.48}$$

With the help of this equation we finally derive

$$\frac{d\mu}{dt} = \frac{d}{dt}\left(\frac{mv_\perp^2}{2B}\right) = \frac{m}{2B}\frac{dv_\perp^2}{dt} - \frac{mv_\perp^2}{2B^2}\frac{dB}{dt} = \frac{1}{B}\left(\frac{m}{2}\frac{dv_\perp^2}{dt} - \mu\frac{dB}{dt}\right) = 0\ . \tag{3.49}$$

From the constancy of μ the possibility of a reflection at the magnetic mirror is readily seen: As B increases, so must $v_\perp^2$, and this is at the expense of $v_\parallel^2$ because the kinetic energy of the particle is fixed; thus $v_\parallel$ may reach zero and reverse its sign. On the other hand, particles with very large $v_\parallel^2$ may penetrate the bottleneck, and so get lost from the magnetic bottle.

Another result can easily be derived from the adiabatic invariant μ, namely that the magnetic flux enclosed by the gyrating particle is constant:

$$\Phi = \pi r_{\mathrm{L}}^2 B = \pi\frac{v_\perp^2 m^2}{q^2 B} = \frac{2\pi m}{q^2}\,\mu\ . \tag{3.50}$$

Therefore, the helix described by the gyrating particle fits into a "mantle" of lines of force.

The name "adiabatic invariant" is derived from the fact that, according to the approximations made, the particle experiences only a weak (or slow) change of the field as it gyrates around the lines of force. The term is also used in thermodynamics where a slow process carries a system through a sequence of equilibrium states (in fact "quasi-static" would be the more appropriate adjective).

3.6.2 Two more adiabatic invariants

In a mechanical system that performs a periodic variation, with period T, with respect to two generalized coordinates p and q the action integral over a full period

$$I = \oint p\,dq \tag{3.51}$$

is invariant against a slow change of a parameter, say λ, on which the system depends (see, e.g., Landau & Lifschitz, Theoretical Physics, Vol. I). The change must be slow in the sense $|Td\lambda/dt| \ll |\lambda|$. The integral I is called an adiabatic invariant.

For a particle that gyrates in a magnetic field the slowly changing parameter is the field strength; as seen from the particle on its path, the field strength varies even in a time-independent field, due to its variation in space. The condition, then, is that the Larmor radius is small in comparison to the characteristic length L of the field inhomogeneity. If the field does change with time, then the characteristic time for that change must be long in comparison to the gyration period, $2\pi/\omega_{\mathrm{L}}$.

For a charged particle in the ionosphere there are three periodicities: First, the gyration around the lines of force, with $p_1 = mv_\perp$ and q_1 the coordinate along

the gyration circle; second, the North-South motion in form of a helix inside the magnetic bottle, with $p_2 = mv_\parallel$ and q_2 a coordinate along the magnetic field; and third, the azimuthal drift (grad$B$ + curvature), with $p_3 = v_\mathrm{D}$ and q_3 the coordinate along the path around the Earth. Hence

$$I_1 = m \oint v_\perp \, dq_1 \, , \tag{3.52}$$

$$I_2 = m \oint v_\parallel \, dq_2 \, , \tag{3.53}$$

$$I_3 = m \oint v_\mathrm{D} \, dq_1 \, , \tag{3.54}$$

are three adiabatic invariants (aside from a constant factor, I_1 is the magnetic moment μ). The integrals are taken over a full period of the three motions, respectively.

Chapter 4

Numerical Methods II – partial differential equations

In this chapter, we will consider partial differential equations of the type

$$\partial_t y = F(y, \partial_x, \partial_x^2 y; x, t) \tag{4.1}$$

and try to approximate the solution $y = y(x,t)$. After discretising in x and t, the values we are interested in are

$$y_l^{(n)} \equiv y(x_l, t_n) \ , \tag{4.2}$$

where

$$x_l = x_0 + l\delta x \ , \tag{4.3}$$

i. e. we use an equidistant grid in x. We are looking for *explicit finite difference schemes* which give us a rule for constructing y_l at a new time step from the values at the previous step,

$$y_l^{(n)} \longmapsto y_l^{(n+1)} \tag{4.4}$$

(we are interested in single-step methods only).

We will use the advection equation

$$\partial_t y = -u\,\partial_x y \ . \tag{4.5}$$

with constant advection velocity u as a sample equation. The exact solution of Eq. (4.5) is given by

$$y(x,t) = y(x-ut, 0) \ . \tag{4.6}$$

Equation (4.5) is the simplest example of a *hyperbolic equation*, i. e. a transport equation with finite transport speed. This is the class of equations one is primarily interested in when doing hydrodynamics. We will often add a small diffusive term to the right-hand side of hyperbolic equations, turning them formally into parabolic equations; nevertheless, our major interest is in advection and wave propagation, and it is reasonable to consider the problems we are dealing with as hyperbolic.

Most of the results derived or outlined here for the simple advection equation (4.5) still hold for other hyperbolic problems like the propagation of weak sound waves.

4.1 A low-order scheme

Approximate ∂_x and ∂_x^2 by central finite differences. To second order in δx, this yields

$$\partial_x y_l = \frac{y_{l+1} - y_{l-1}}{2\delta x} + O\left(\delta x^2\right) , \tag{4.7}$$

$$\partial_x^2 y_l = \frac{y_{l+1} - 2y_l + y_{l-1}}{\delta x^2} + O\left(\delta x^2\right) . \tag{4.8}$$

$$\tag{4.9}$$

Note that we use only values on one single mesh in x — we are using a 'non-staggered mesh', which makes many things simpler. Applying this discretisation to the right-hand side of Eq. (4.5) and applying first-order (Euler) discretisation in time as in § 1.2, we can write

$$\frac{y_l^{(n+1)} - y_l^{(n)}}{\delta t} = u\, \frac{y_{l+1}^{(n)} - y_{l-1}^{(n)}}{2\delta x} , \tag{4.10}$$

which yields an explicit formula for $y_l^{(n+1)}$,

$$y_l^{(n+1)} = y_l^{(n)} - \frac{u\,\delta t}{\delta x}\, \frac{y_{l+1}^{(n)} - y_{l-1}^{(n)}}{2} . \tag{4.11}$$

Press et al. (1996) state:

> *"The resulting finite-difference approximation [...] is called the FTCS representation (Forward Time Centred Space) [...] It's a fine example of an algorithm that is easy to derive, takes little storage, and executes quickly. Too bad it doesn't work!"*

Figure 4.1 shows how small scale structures eventually destroy the initial pattern in the case of the advection equation (4.5), even for a very small time step of $\delta t = 0.0005$. The perturbations grow faster if the time step is larger; obviously, this scheme is *unstable*.

We can try three things to stabilise the scheme:

1. use higher spatial order;
2. use higher time order;
3. add viscosity.

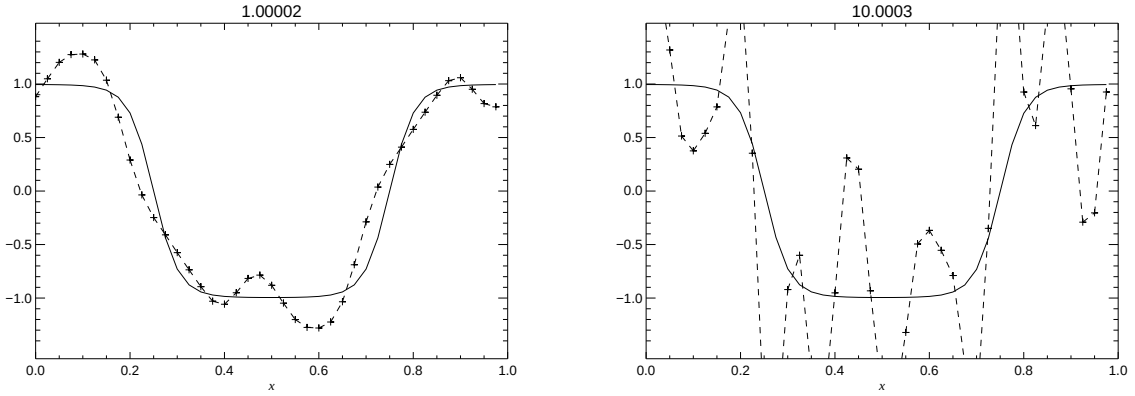

Figure 4.1: The $O\left(\delta t, \delta x^2\right)$ (first-order time step, second-order spatial derivatives) scheme applied to the advection problem (4.5) with $u = 1$ and periodic boundary conditions. The time step is extremely small ($\delta t = 0.0005$). The solid line shows the exact solution (identical to the initial profile), while the crosses and dashed line show the numerical solution. Left: t=1 (i.e. the pattern has travelled once through the interval $[0, 1]$. Right: t=10 (the pattern has travelled ten times through the interval).

4.2 Higher-order schemes

4.2.1 Higher spatial order

Ansatz for nth order finite difference *stencil*:

$$y'_0 = \sum_{p=-n/2}^{n/2} w_p y_p + O\left(\delta x^n\right) \equiv D^{(1)} y_0 + O\left(\delta x^n\right) , \tag{4.12}$$

$$y''_0 = \sum_{p=-n/2}^{n/2} \widetilde{w}_p y_p + O\left(\delta x^n\right) \equiv D^{(2)} y_0 + O\left(\delta x^n\right) , \tag{4.13}$$

where $n = 2, 4, 6, \ldots$ Figure 4.2 gives a schematic view of the information used for a fourth-order (five-point) finite-difference stencil.

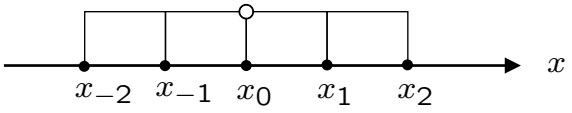

Figure 4.2: Sketch of a 5-point finite-difference stencil.

We determine the coefficients w_p, $\widetilde{w}_p$ from the requirement that the exact values of the derivatives are retrieved for all polynomials of up to nth order, and in particular for the monomials $y = 1$, $y = x$, $y = x^2$, …, $y = x^n$ at $x_0 = 0$. As an example, we derive the conditions for the 4th-order (5-point) first derivative

Table 4.1: Coefficients $w_k = p_k/q = -w_{-k}$ and residual term R for finite difference approximations of order n to the first derivative, $y'_0 = \sum_k w_k y_k + R$. The point ξ is in the interval $x_{-n/2} < \xi < x_{n/2}$.

n	p_0	p_1	p_2	p_3	p_4	p_5	q	R
2	0	1					$2\delta x$	$-\frac{1}{6}y^{(3)}(\xi)\delta x^2$
4	0	8	-1				$12\delta x$	$\frac{1}{30}y^{(5)}(\xi)\delta x^4$
6	0	45	-9	1			$60\delta x$	$-\frac{1}{140}y^{(7)}(\xi)\delta x^6$
8	0	672	-168	32	-3		$840\delta x$	$+\frac{1}{630}y^{(9)}(\xi)\delta x^8$
10	0	2100	-600	150	-25	2	$2520\delta x$	$-\frac{1}{2772}y^{(11)}(\xi)\delta x^{10}$

operator:

$$\begin{array}{rrrrrrrrrrr} w_{-2} + & w_{-1} + & w_0 + & w_1 + & w_2 & = & 0\,, \\ -2\,\delta x\, w_{-2} - & \delta x\, w_{-1} & & + \delta x\, w_1 + & 2\,\delta x\, w_2 & = & 1\,, \\ 4\,\delta x^2 w_{-2} - & \delta x^2 w_{-1} & & + \delta x^2 w_1 + & 4\,\delta x^2 w_2 & = & 0\,, \\ & & \cdots & & & & \cdots \end{array} \tag{4.14}$$

i. e.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ (-2)^2 & (-1)^2 & 0 & 1^2 & 2^2 \\ (-2)^3 & (-1)^3 & 0 & 1^3 & 2^3 \\ (-2)^4 & (-1)^4 & 0 & 1^4 & 2^4 \end{pmatrix} \begin{pmatrix} w_{-2} \\ w_{-1} \\ w_0 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1!/\delta x \\ 0 \\ 0 \\ 0 \end{pmatrix} . \tag{4.15}$$

The system for the coefficients $\widetilde{w}_p$ of the second-derivative operator differs only in the right-hand side:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ (-2)^2 & (-1)^2 & 0 & 1^2 & 2^2 \\ (-2)^3 & (-1)^3 & 0 & 1^3 & 2^3 \\ (-2)^4 & (-1)^4 & 0 & 1^4 & 2^4 \end{pmatrix} \begin{pmatrix} \widetilde{w}_{-2} \\ \widetilde{w}_{-1} \\ \widetilde{w}_0 \\ \widetilde{w}_1 \\ \widetilde{w}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2!/\delta x^2 \\ 0 \\ 0 \end{pmatrix} . \tag{4.16}$$

Tables 4.1 and 4.2 list the coefficients w_p and $\widetilde{w}_p$ for schemes of up to tenth order. For convenience we write out explicitly the sixth-order formulas:

$$y'_0 = \frac{-y_{-3} + 9y_{-2} - 45y_{-1} + 45y_1 - 9y_2 + y_3}{60\,\delta x} + O\left(\delta x^6\right) \tag{4.17}$$

$$y''_0 = \frac{2y_{-3} - 27y_{-2} + 270y_{-1} - 490y_0 + 270y_1 - 27y_2 + 2y_3}{180\,\delta x^2} + O\left(\delta x^6\right) . \tag{4.18}$$

Table 4.2: Coefficients $\widetilde{w}_k = p_k/q = \widetilde{w}_{-k}$ and residual term R for finite difference approximations of order n to the second derivative, $y_0'' = \sum_k \widetilde{w}_k y_k + R$. The point ξ is in the interval $x_{-n/2} < \xi < x_{n/2}$.

n	p_0	p_1	p_2	p_3	p_4	p_5	q	R
2	-2	1					δx^2	$-\frac{1}{12}y^{(4)}(\xi)\delta x^2$
4	-30	16	-1				$12\delta x^2$	$\frac{1}{90}y^{(6)}(\xi)\delta x^4$
6	-490	270	-27	2			$180\delta x^2$	$-\frac{1}{560}y^{(8)}(\xi)\delta x^6$
8	-14350	8064	-1008	128	-9		$5040\delta x^2$	$\frac{1}{3150}y^{(10)}(\xi)\delta x^8$
10	-73766	42000	-6000	1000	-125	8	$25200\delta x^2$	$-\frac{1}{16632}y^{(12)}(\xi)\delta x^{10}$

4.2.2 Spectral characteristics of finite-difference stencils

Important insight into the properties of a finite-difference scheme is obtained by applying it to a harmonic function

$$y = e^{ikx} \ . \tag{4.19}$$

Here $0 \le |k| \le \pi/\delta x$, because the *Nyquist wave number*

$$k_{\rm Ny} \equiv \frac{\pi}{\delta x} \tag{4.20}$$

is the highest wave number that can be distinguished on a grid of spacing δx.[1]

Applying the exact first and second derivative operators to the harmonic function (4.19) would yield

$$\partial_x e^{ikx} = ike^{ikx} \ , \qquad \partial_x^2 e^{ikx} = -k^2 e^{ikx} \ , \tag{4.22}$$

thus the *spectral transfer functions*

$$H^{(1)}(k) \equiv e^{-ikx} D^{(1)} e^{ikx} \ , \tag{4.23}$$

$$H^{(2)}(k) \equiv e^{-ikx} D^{(2)} e^{ikx} \tag{4.24}$$

indicate the quality of the finite-difference approximations $D^{(1)}$, $D^{(2)}$: for exact derivatives one would get [2]

$$H^{(1)}(k) = ik \ , \qquad H^{(2)}(k) = -k^2 \ . \tag{4.25}$$

[1] According to

$$e^{i(k_{\rm Ny}+k')x_l} = (-1)^l e^{ik'x_l} = e^{i(-k_{\rm Ny}+k')x_l} \ , \tag{4.21}$$

the wave number $k_{\rm Ny}+k'$ is equivalent to $-k_{\rm Ny}+k'$.

[2] Such exact numerical derivative operators are indeed implemented by *spectral schemes* which apply a Fourier transform, multiply in Fourier space by ik or $-k^2$, and then transform back.

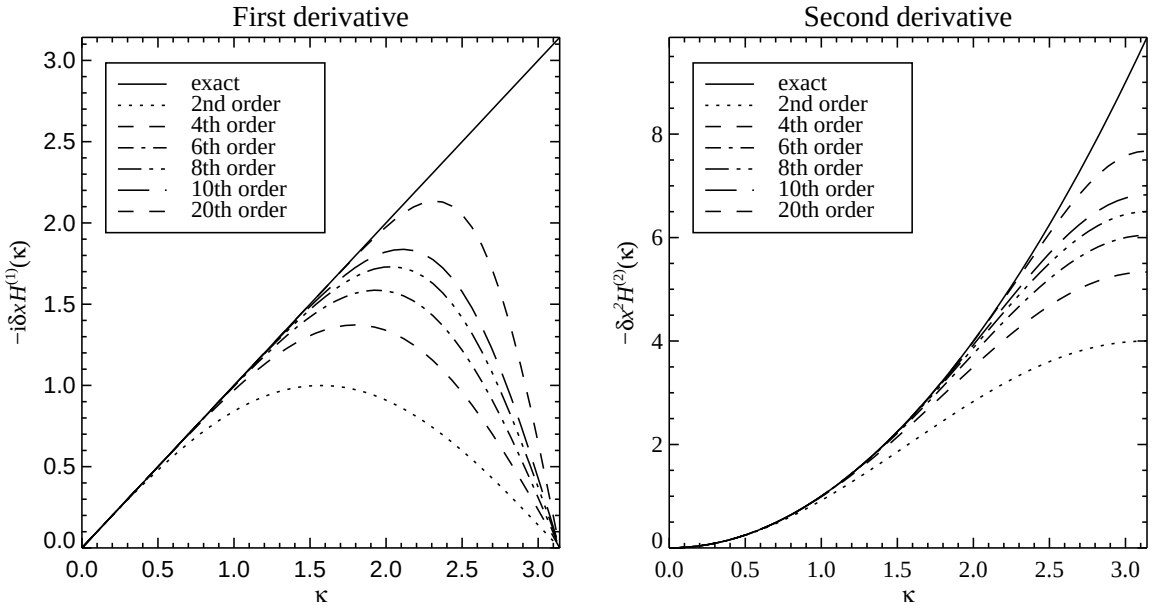

Figure 4.3: Spectral transfer functions $H(k\delta x) \equiv e^{-ik} D e^{ik}$ as a function of $\kappa = k\delta x$ for centred finite-difference schemes of different orders. Left: transfer function for the first derivative operator, $D^{(1)}$, multiplied by $-i\,\delta x$. Right: spectral transfer function for the second derivative operator, $D^{(2)}$, multiplied by $-\delta x^2$. The solid lines show the transfer function of the exact derivative operator (which is reproduced by spectral schemes).

Figure 4.3 shows the spectral transfer functions for a number of schemes from order 2 up to 20. One can easily see how all schemes yield good approximations to the exact derivative for small k, but for intermediate wave numbers (say, half the Nyquist wavenumber $\kappa_{\mathrm{Ny}} = k_{\mathrm{Ny}}\delta x = \pi$) only higher orders reproduce the exact derivatives with sufficient accuracy.

As for stability, it turns out that the numerical solution of our simple advection problem is getting even more unstable (instabilities grow faster) with higher-order spatial derivative operators. So we have to try our next trick.

4.2.3 Higher-order time-stepping schemes

If we apply some of the higher-order time-stepping schemes from Chapter 1, the situation changes drastically. The stability behaviour can be investigated by Fourier-mode stability analysis which we will only outline here. Consider harmonic modes of the form

$$y = \hat{y} e^{\gamma_* t} e^{ikx} \; . \tag{4.26}$$

We know that for the advection equation (4.5) — and for some other important linear test problems — the energy in any Fourier mode remains the same, so ideally γ_* should just have an imaginary part; for the exact solution to Eq. (4.5), it would be $\gamma_* = -iuk$. The numerical scheme will still have solutions of the form

(4.26), but now γ_* will deviate from the exact result,

$$\gamma_* = -iuk + \gamma + i\omega \; . \tag{4.27}$$

The quantity γ is a growth rate and gives rise to the *amplitude error*, while ω represents the *phase error*.

The question of stability boils down to the sign of γ. If $\gamma > 0$ for some modes, then the energy in these modes will grow and eventually dominate the solution and render it useless. Reducing the *modulus* of γ in this case will not remove the instability — it only increases the time for which it can be be ignored.

If $\gamma < 0$, on the other hand, energy in the corresponding modes will decrease. This implies that there is some *numerical dissipation* at work, but normally this only affects the smaller scales. By decreasing the time step, both amplitude and phase error will be decreased, so if $\gamma \leq 0$ for all modes, one can control the errors by adjusting the time step δt.

Table 4.3: Leading-order terms (in δt) of the growth rate γ and phase drift ω for time-stepping schemes of different order m. The quantity $0 < \Theta \equiv H^{(1)}/(ik) < 1$ is a measure of the quality of the spatial scheme: $\Theta \approx 1$ where the scheme works good (typically for small k). Note that $\gamma < 0$ (indicating stability) only for $m = 3, 4; 7, 8; 11, 12 \ldots$.

m	γ	ω
1	$\frac{(uk\Theta)^2}{2}\delta t$	$uk(1-\Theta) + \frac{(uk\Theta)^3}{3}\delta t^2$
2	$\frac{(uk\Theta)^4}{8}\delta t^3$	$uk(1-\Theta) - \frac{(uk\Theta)^3}{6}\delta t^2$
3	$-\frac{(uk\Theta)^4}{24}\delta t^3$	$uk(1-\Theta) - \frac{(uk\Theta)^5}{30}\delta t^4$
4	$-\frac{(uk\Theta)^6}{144}\delta t^5$	$uk(1-\Theta) + \frac{(uk\Theta)^5}{120}\delta t^4$

Table 4.3 shows the leading order in δt of the amplitude and phase errors for time-stepping schemes of orders 1 to 4. Only the third- and fourth-order schemes are stable ($\gamma < 0$), provided that the time step is sufficiently small (see § 4.4 below). Although this result is formulated for the advection problem (4.5), exactly the same stability conditions hold in the case of linear sound waves

$$\partial_t \ln \varrho = -\partial_x v \tag{4.28}$$

$$\partial_t v = -c_\mathrm{s}^2 \partial_x \ln \varrho \tag{4.29}$$

if the advection speed u is replaced by the speed of sound c_s. In the case of sound waves in a medium that moves at speed u, the relevant velocity is $\max(|u \pm c_\mathrm{s}|)$.

To conclude, we can say that (for advection and similar problems) the amplitude error, and thus the stability of the scheme, is determined by the time-stepping scheme, while the phase error is normally dominated by the spatial discretisation.

4.3 Artificial viscosity

A look at Table 4.3 shows that for Euler time stepping the growth rate $\gamma = (uk\Theta)^2\delta t/2$ is proportional to k^2 and thus corresponds to a negative *numerical diffusivity* (often called *numerical viscosity*), because a positive diffusivity ν would give rise to a negative growth rate $\gamma = -\nu k^2$. From this we see that we can compensate the growth of the modes by adding an additional viscous term to the right-hand side of Eq. (4.5),

$$\partial_t y = -u\,\partial_x y + \nu\partial_x^2 y\ . \tag{4.30}$$

If we choose

$$\nu = \frac{1}{2}u^2\,\delta t\ , \tag{4.31}$$

the scheme becomes stable. A diffusivity ν added to stabilise a scheme is often referred to as *artificial viscosity*. The scheme obtained by discretising Equ. (4.30) is will be stable, provided the time step is small enough. Note that if $\delta x \to 0$, δt will also have to go to zero (see § 4.4 below); then, according to (4.31), the artificial viscosity ν will tend to zero in this limit, too. This implies that the solution we approach with ever finer meshes will be the solution of the original equation (4.5) without any diffusion term.[3]

When solving partial differential equations that are more realistic than our simple advection problem, even third- and fourth order time-stepping schemes require some amount of diffusivity/viscosity due to boundary effects, nonlinearities or just to minimise the consequences of the phase error. Like in the case discussed above, this viscosity will always tend to zero for $\delta x \to 0$. The recommended minimum value of viscosity for the $O\left(\delta t^3, \delta x^6\right)$ scheme is

$$\nu = c_\nu\, U_{\max}\, \delta x \tag{4.32}$$

where $U_{\max}$ is the largest velocity in the problem (including propagation speeds of waves), and $c_\nu = 0.01\,..\,0.02$.

4.4 The length of the time step

Even the explicit schemes labelled as 'stable' are only stable if the time step δt satisfies the *Courant condition*

$$\delta t \le c_{\mathrm{adv}}\,\frac{\delta x}{u}\ , \tag{4.33}$$

where c_{adv} is a dimensionless number of order unity.[4] The ratio

$$\mathcal{C} \equiv \frac{u\delta t}{\delta x} \tag{4.34}$$

[3] This property is referred to as *consistency* of a scheme. One popular example of a scheme that can be inconsistent is the *Dufort–Frankel scheme*.

[4] This can be different for *implicit schemes* where one needs to solve a system of equations to obtain $y_l^{(n+1)}$ from $y_l^{(n)}$.

is called the *Courant number* or *CFL number* (after Courant, Friedrichs and Levy). For $O(\delta t^3, \delta x^6)$ schemes, the stability boundary is $c_{\text{adv}} = 1.092$. For $O(\delta t^4, \delta x^6)$ schemes, we have $c_{\text{adv}} = 1.783$. In practise one should use a time step considerably smaller than the stability limit ($\mathcal{C} = 0.5$ or smaller), since at the very limit the accuracy will always be poor.

If the equation to be solved contains (physical or artificial) diffusive terms, there is another stability limit to the time step,

$$\delta t \leq c_{\text{diff}} \frac{\delta x^2}{\nu} \ . \tag{4.35}$$

In fact, this is simply another manifestation of condition (4.33) if we introduce the velocity $U_{\text{diff}} = \nu/\delta x$ associated with diffusion. Reasonable values for the coefficients c_{adv} and c_{diff} can however be quite different.

Other propagation velocities (like the sound speed) will give rise to similar time step restrictions. The recommended time step for the $O(\delta t^3, \delta x^6)$ scheme is

$$\delta t = \min\left(0.4\, U_{\max}\delta x,\ 0.08\, \nu_{\max}\delta x^2\right) \ , \tag{4.36}$$

where $U_{\max}$ is the largest velocity in the problem and $\nu_{\max}$ the largest diffusivity.

Interpretation of the Courant criterion As a rule of thumb, the time step restrictions (4.33) and (4.35) can be interpreted as follows: explicit finite-difference schemes are only stable if the time step is small enough for information to traverse essentially only one grid cell per time step. However, this is not a strict statement at all. One can easily see that it is the phase speed, rather than the group speed of a given wave family that determines the stability of a scheme, thus the issue is not really one of propagation of information.

Our standard scheme

This box summarises the properties of the scheme we normally use to solve partial differential equations.

- 6th-order spatial derivative operators (4.17), (4.18);
- 3rd-order Runge–Kutta timestepping (see page 1.3);
- Artificial viscosity:
$$\nu = c_\nu U_{\max}\delta x$$
with $c_\nu = 0.01 \ldots 0.02$;
- Time step:
$$\delta t = \min\left(0.4\, U_{\max}\delta x,\ 0.08\, \nu_{\max}\delta x^2\right) \ .$$

4.5 Boundary conditions

So far, our discussion was implicitly assuming that boundary conditions are periodic (or that the interval in x is unbounded). In real life, one often has to use other boundary conditions. We will discuss this just briefly, restricting ourselves to boundary conditions implemented by setting *ghost zone* values. A ghost zone is a layer of fictitious points beyond the boundary which is introduced so that wide finite-difference stencils can be applied even close to the boundary. For our sixth-order (seven-point) stencil, we need three points on each side of the given point, thus we will need three ghost layers on each side if we want to be able to calculate derivatives in the very boundary points. This situation is depicted in Fig. 4.4.

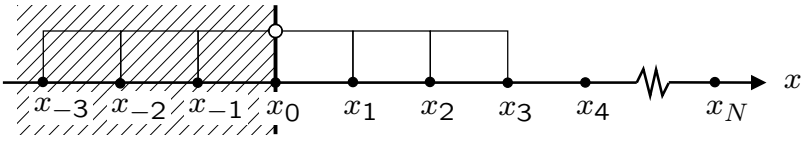

Figure 4.4: Sketch of ghost zones for a seven-point finite-difference stencil on a grid ranging from x_0 to x_N.

When ghost zones are used, the boundary conditions provide a rule how to set the values in the ghost points. We just present four popular choices of boundary conditions that can be thus implemented:

1. Periodic boundary conditions

$$y_{-1} = y_{N-1} \ , \qquad y_{-2} = y_{N-2} \ , \qquad y_{-3} = y_{N-3} \ . \tag{4.37}$$

2. Symmetry ($y_0' = 0$)

$$y_{-1} = y_1 \ , \qquad y_{-2} = y_2 \ , \qquad y_{-3} = y_3 \ . \tag{4.38}$$

3. Antisymmetry ($y_0 = 0$)

$$y_{-1} = -y_1 \ , \qquad y_{-2} = -y_2 \ , \qquad y_{-3} = -y_3 \ . \tag{4.39}$$

4. Generalised antisymmetry ($y_0'' = 0$)

$$y_{-1} = 2y_0 - y_1 \ , \qquad y_{-2} = 2y_0 - y_2 \ , \qquad y_{-3} = 2y_0 - y_3 \ . \tag{4.40}$$

Figure 4.5 illustrates these four boundary conditions

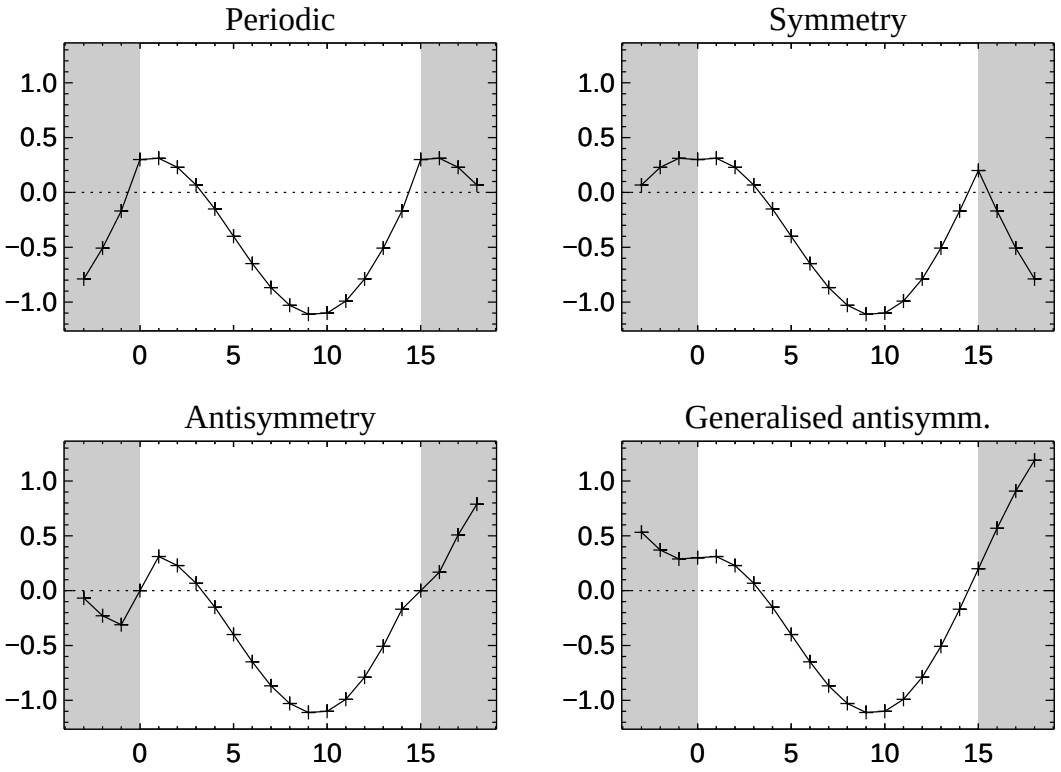

Figure 4.5: The four boundary conditions discussed in the text, applied to an arbitrarily chosen function. The shaded regions to the left and right are the ‘ghost zones’.

Chapter 5

Stellar Winds and Critical Points

5.1 Fluid dynamics

Fluid dynamics or *hydrodynamics* is governed by conservation of mass, momentum and possibly other quantities.

5.1.1 Mass conservation

Conservation of mass requires that the mass $M = \int_{\mathcal{V}} \varrho \; dV$ in a given (*fixed*) volume $\mathcal{V}$, changes according due to the mass flux F_m through the surface, $F_m = \int_{\partial\mathcal{V}} \mathbf{j}_m \cdot \, d\mathbf{f}$, where ϱ denotes (mass) density, $\mathbf{j}_m = \varrho \mathbf{v}$ the mass flux density, and $\mathbf{v}$ the velocity of the fluid (see Fig. 5.1 for an illustration). Thus, our mass balance has the form

$$\frac{dM}{dt} \; = \; \frac{d}{dt} \int\limits_{\mathcal{V}} \varrho \; dV = \; - \int\limits_{\partial\mathcal{V}} \varrho \mathbf{v} \cdot \, d\mathbf{f} \quad . \tag{5.1}$$

Pulling the time derivative into the integrand on the left-hand side and applying Gauß' integral theorem to the right-hand side, we can write

$$\int\limits_{\mathcal{V}} \left[\frac{\partial \varrho}{\partial t} + \mathrm{div}(\varrho \mathbf{v}) \right] \, dV = 0 \; . \tag{5.2}$$

This must hold for any fixed volume $\mathcal{V}$; thus if we contract $\mathcal{V}$ onto a given point $\mathbf{x}$, we obtain the *continuity equation*

$$\frac{\partial \varrho}{\partial t} + \mathrm{div}(\varrho \mathbf{v}) = 0 \; . \tag{5.3}$$

5.1.2 Momentum conservation

The continuity equation (5.3) can also be written in index notation,

$$\frac{\partial \varrho}{\partial t} + \frac{\partial j_k}{\partial x_k} = 0 \; , \tag{5.4}$$

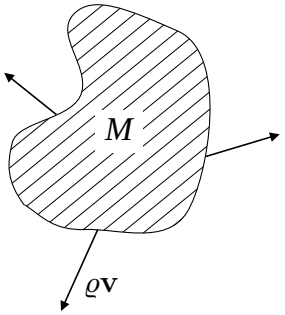

Figure 5.1: Sketch of the mass balance in a volume $\mathcal{V}$. The total mass M can only change in agreement with the mass flux $\varrho\mathbf{v}$ through the boundary.

with $j_k \equiv \varrho v_k$ denoting the mass flux density as above. In direct analogy, conservation of momentum is expressed by a continuity equation of the form

$$\frac{\partial \varrho v_i}{\partial t} + \frac{\partial \Pi_{ik}}{\partial x_k} = 0 \ , \tag{5.5}$$

where ϱv_i is momentum density and Π_{ik} the momentum flux density, which is a symmetric tensor describing the flux of i-momentum in the k-direction. The tensor Π_{ik} consists of a transport component $\varrho v_i v_k$ (reminiscent of the mass flux density above) and the stress tensor σ_{ik}. The latter can itself be split into the isotropic pressure tensor $p\delta_{ik}$ and the viscous stress $-\sigma_{ik}^{(visc)}$ (the minus sign is convention). Thus,

$$\Pi_{ik} = \varrho v_i v_k + p\delta_{ik} - \sigma_{ik}^{(visc)} \ , \tag{5.6}$$

and Eq. (5.5) takes the form

$$\frac{\partial(\varrho v_i)}{\partial t} + \frac{\partial}{\partial x_k}\left[\varrho v_i v_k + p\delta_{ik} - \sigma_{ik}^{(visc)}\right] = 0 \ . \tag{5.7}$$

This can be expanded to

$$\varrho\partial_t v_i + v_i\partial_t\varrho + v_i\partial_k(\varrho v_k) + \varrho v_k\partial_k v_i + \partial_i p - \partial_k\sigma_{ik}^{(visc)} = 0 \ . \tag{5.8}$$

The second and third term together are equal to $v_i\left[\partial_t\varrho + \mathrm{div}(\varrho\mathbf{v})\right]$, which vanishes because of the continuity equation (5.3), and thus after dividing by ϱ we are left with

$$\partial_t v_i + v_k\partial_k v_i = -\frac{1}{\varrho}\partial_i p + \frac{1}{\varrho}\partial_k\sigma_{ik}^{(visc)} \ . \tag{5.9}$$

The viscous stress tensor $\sigma_{ik}^{(visc)}$ must be symmetric. If it is assumed to be linear in $\partial_i u_k$ and traceless, it must have the form

$$\sigma_{ik}^{(visc)} = \mu\left(\partial_i v_k + \partial_k v_i - \frac{2}{3}\,\mathrm{div}\,\mathbf{v}\,\delta_{ik}\right) \ , \tag{5.10}$$

where the coefficient μ is called *dynamical viscosity*. Assuming $\mu = \text{const}$, we can write

$$\partial_k\sigma_{ik}^{(visc)} = \mu\left(\partial_i\partial_k v_k + \partial_k^2 v_i - \frac{2}{3}\partial_i\partial_l u_l\right) = \mu\left(\Delta\, v_i + \frac{1}{3}\partial_i\,\mathrm{div}\,\mathbf{v}\right) \ ; \tag{5.11}$$

thus, the *equation of motion* is

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \operatorname{grad})\mathbf{v} = -\frac{1}{\varrho} \operatorname{grad} p + \nu \left(\Delta \mathbf{v} + \frac{1}{3} \operatorname{grad} \operatorname{div} \mathbf{v} \right) + \mathbf{f}_{\text{ext}} \ , \tag{5.12}$$

where $\nu \equiv \mu/\varrho$ is the *kinematic viscosity*, and we have added the acceleration $\mathbf{f}_{\text{ext}}$ due to external forces like gravity. Equation (5.12) is often referred to as the *Navier–Stokes equation*. In the one-dimensional case, $\partial_y{=}\partial_z{=}0$, $\mathbf{v}{=}(u,0,0)$, it reduces to

$$\partial_t u + u \partial_x u = -\frac{1}{\varrho} \partial_x p + \frac{4}{3} \nu \, \partial_x^2 u \ . \tag{5.13}$$

5.1.3 The pressure term

If there is a unique relation between p and ϱ — for example an adiabatic, polytropic or isothermal equation of state — we can define the sound speed [1]

$$c_{\text{s}}^2 \equiv \frac{dp}{d\varrho} \ . \tag{5.16}$$

This allows us to rewrite the pressure term as follows

$$-\frac{1}{\varrho} \operatorname{grad} p = -\frac{c_{\text{s}}^2}{\varrho} \operatorname{grad} \varrho = -c_{\text{s}}^2 \operatorname{grad} \ln \varrho \ . \tag{5.17}$$

As $\ln \varrho$ turns out to be a useful variable, we divide the continuity equation (5.3) by ϱ and express it in terms of $\ln \varrho$, too:

$$\frac{\partial_t \varrho}{\varrho} + \frac{\mathbf{v} \cdot \operatorname{grad} \varrho + \varrho \operatorname{div} \mathbf{v}}{\varrho} = \partial_t \ln \varrho + \mathbf{v} \cdot \operatorname{grad} \ln \varrho + \operatorname{div} \mathbf{v} = 0 \tag{5.18}$$

[1] For a perfect gas in the adiabatic case (entropy $s = \text{const}$), the equation of state is

$$p = K \varrho^{\gamma} \ , \tag{5.14}$$

where $\gamma \equiv c_p/c_v$ is the adiabatic index, i. e. the ratio of specific heat at constant pressure, c_p, to the specific heat at constant volume, c_v, and K is a constant related to the entropy s. For this case we obtain the familiar relation

$$c_{\text{s}}^2 = \left(\frac{dp}{d\varrho} \right)_s = \gamma \frac{p}{\varrho} = \gamma \frac{\mathcal{R}}{\mu_{\text{mol}}} T \tag{5.15}$$

where $\mathcal{R}/\mu_{\text{mol}}$ is the specific gas constant, T the temperature, and $(\partial/\partial)_s$ denotes the partial derivative for constant entropy s.

A *polytropic equation of state* looks like the adiabatic one (5.14), but with the adiabatic index replaced by an exponent Γ that is treated as a free parameter.

Thus, in terms of logarithmic density, our equations become

$$\frac{\partial \ln \varrho}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \ln \varrho = -\operatorname{div} \mathbf{v} \tag{5.19}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \operatorname{grad})\mathbf{v} = -c_\mathrm{s}^2 \operatorname{grad} \ln \varrho + \nu \left(\Delta \mathbf{v} + \frac{1}{3} \operatorname{grad} \operatorname{div} \mathbf{v} \right) + \mathbf{f}_\mathrm{ext} \; . \tag{5.20}$$

Sound waves
If we linearise equations (5.19), (5.20) using the ansatz

$$\ln \varrho = \ln \varrho_0 + \lambda \; , \tag{5.21}$$

$$\mathbf{v} = \mathbf{0} + \mathbf{u} \tag{5.22}$$

and assuming that $\lambda \ll 1$, $|\mathbf{u}| \ll c_\mathrm{s}$, we obtain the system

$$\frac{\partial \lambda}{\partial t} = -\frac{\partial u}{\partial x} \; , \tag{5.23}$$

$$\frac{\partial u}{\partial t} = -c_\mathrm{s}^2 \frac{\partial \lambda}{\partial x} \; . \tag{5.24}$$

This system has the general solution

$$\lambda = f(x - c_\mathrm{s} t) + g(x + c_\mathrm{s} t) \tag{5.25}$$

$$u = c_\mathrm{s} \, f(x - c_\mathrm{s} t) - c_\mathrm{s} \, g(x + c_\mathrm{s} t) \tag{5.26}$$

where $f(\cdot)$, $g(\cdot)$ are arbitrary functions.

5.2 Parker wind

The solar wind is a continuous flow of ionised gas from the Sun; Fig. 5.2 shows the angular dependency of its velocity as measured by *Ulysses*. Typical velocities and particle densities at the position of the Earth are around $400\,\mathrm{km/s}$ ('fast solar wind') and $10^4\,\mathrm{m}^{-3}$, i. e. 5 protons and 5 electrons per cm^3. Following Parker (1958), we now develop a simple hydrodynamical model for the acceleration of such a wind.

To make the model as simple as possible, we assume the medium to be a perfect gas (a good approximation) and to be isothermal (a bad approximation, but qualitatively still OK). The geometry is assumed to be spherically symmetric, thus the only spatial coordinate is spherical radius r, and the velocity vector has only a radial component, $\mathbf{v} = (v_r, 0, 0)$. For this model, the continuity and Navier–Stokes equations take the form

$$\partial_t \varrho = -\operatorname{div}(\varrho \mathbf{v}) = -\frac{1}{r^2} \partial_r (r^2 \varrho v_r) \; , \tag{5.27}$$

$$\partial_t v_r = -v_r \partial_r v_r - c_\mathrm{s}^2 \, \partial_r \ln \varrho - \frac{GM_\odot}{r^2} \; . \tag{5.28}$$

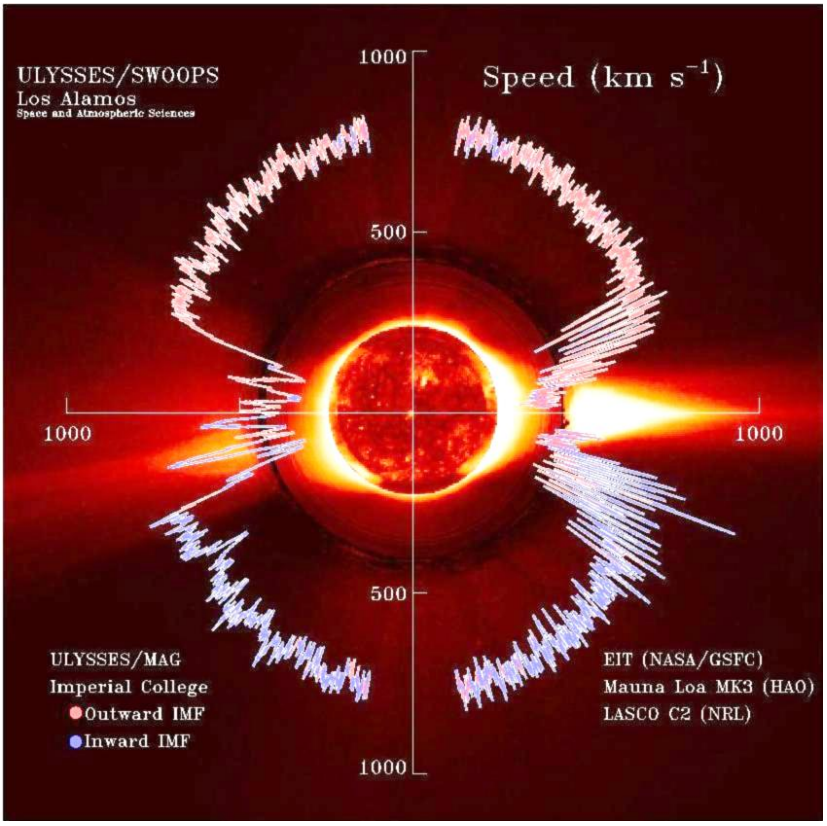

Figure 5.2: Solar wind velocities as measured by the *Ulysses* spacecraft. The background image is a combination of coronographic images showing that the Sun is surrounded by a medium that scatters light (mostly the electrons of the solar wind and similar flows).

Now we assume a steady state, $\partial_t{=}0$, and find that the continuity equation yields

$$\partial_r(r^2\varrho v_r) = 0 \;, \tag{5.29}$$

i. e. the *mass flux*

$$\Phi_\mathrm{m} \equiv 4\pi\varrho v_r r^2 = \text{const} \;. \tag{5.30}$$

We can derive a second invariant from the equation of motion (5.28), which can be written as

$$-\partial_r\left[\frac{v_r^2}{2} + c_\mathrm{s}^2\ln\varrho - \frac{GM_\odot}{r}\right] = 0 \;. \tag{5.31}$$

This implies that

$$\mathcal{E} \equiv \frac{v_r^2}{2} + c_\mathrm{s}^2\ln\varrho - \frac{GM_\odot}{r} = \text{const} \;. \tag{5.32}$$

$\mathcal{E}$ is called the *Bernoulli constant* [2] and should *not* be interpreted as a specific energy — it is only the energy flux density divided by ϱv_r. This is different from the specific energy, which is generally *not* constant.

[2] In the more general case of an adiabatic gas, the Bernoulli constant would be

$$\mathcal{E} \equiv \frac{v_r^2}{2} + h - \frac{GM_\odot}{r} \;, \tag{5.33}$$

The mass flux Φ_m allows us to eliminate ϱ from the equation of motion (5.28): with

$$\ln \varrho = \ln \frac{\Phi_m}{4\pi} - \ln v_r - 2 \ln r \tag{5.35}$$

we can write the pressure term as

$$-c_\mathrm{s}^2 \, \partial_r \ln \varrho = c_\mathrm{s}^2 \left(\partial_r \ln v_r + \frac{2c_\mathrm{s}^2}{r} \right) \ . \tag{5.36}$$

Thus, (5.28) becomes

$$-v_r \partial_r v_r + c_\mathrm{s}^2 \partial_r \ln v_r + \frac{2c_\mathrm{s}^2}{r} - \frac{GM_\odot}{r^2} = 0 \ , \tag{5.37}$$

or

$$(c_\mathrm{s}^2 - v_r^2) \frac{\partial_r v_r}{v_r} = \frac{2c_\mathrm{s}^2}{r^2} (r_* - r) \ , \tag{5.38}$$

where

$$r_* \equiv \frac{GM_\odot}{2c_\mathrm{s}^2} \ . \tag{5.39}$$

Equation (5.38), which can also be written as

$$\frac{\partial v_r}{\partial r} = \frac{2c_\mathrm{s}^2 \, v_r}{r^2} \, \frac{r_* - r}{c_\mathrm{s}^2 - v_r^2} \ , \tag{5.40}$$

shows that the solution can turn from subsonic ($v_r < c_\mathrm{s}$) to supersonic ($v_r > c_\mathrm{s}$) only where $r_* = r$. This solution which passes through the *sonic point* $r = r_*$, $v_r = c_\mathrm{s}$ is sometimes referred to as *sonic solution*. Such points where the enumerator and denominator of the right-hand side of a differential equation must simultaneously turn zero are called *critical points* of the equation.

Figure 5.3 shows all possible solutions of Parker's problem normalised such that the sonic point is at $r_* = 1$, $c_\mathrm{s} = 1$. If the solar wind starts as a subsonic outflow from the solar surface, it can either pass through the sonic point and become supersonic (one of the solid lines in Fig. 5.3), or it can remain subsonic everywhere (the lower set of dotted lines). In the latter case, the velocity will decrease for large r ('solar breeze'), and one finds that pressure and density would remain finite for $r \to \infty$, and in particular at the Earth's orbit they would be much larger than they are found to be. The only explanation reconciling observations and Parker's theory is that the solar wind indeed follows the sonic solution.

---

where

$$h = \int \frac{dp}{\varrho} = \frac{c_\mathrm{s}^2}{\gamma - 1} \tag{5.34}$$

is the specific *enthalpy* of the gas.

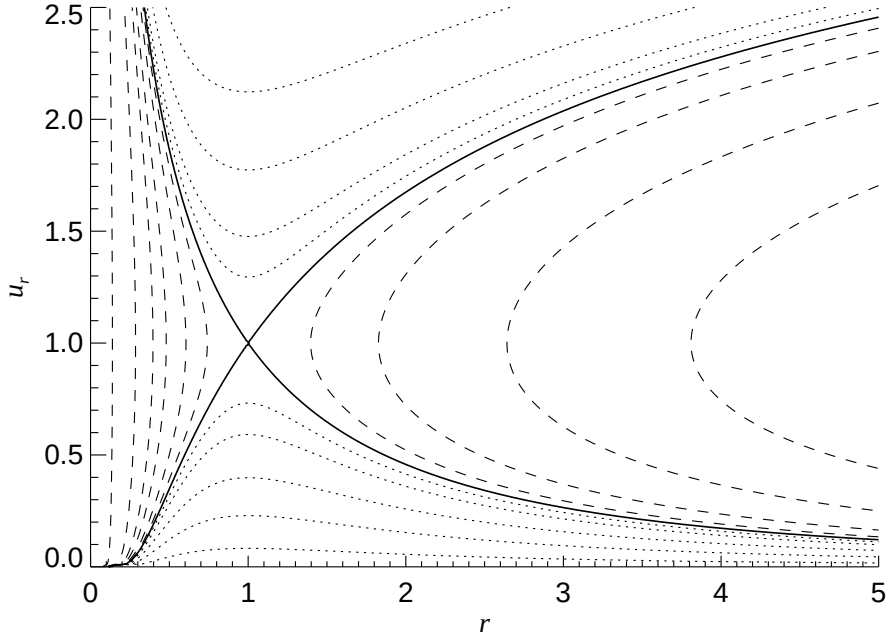

Figure 5.3: Solutions of Parker's problem (5.38) for $c_\mathrm{s} = 1$, $GM_\odot = 2$. Shown are isocontours of $\widetilde{\mathcal{E}} := \mathcal{E} - c_\mathrm{s}^2 \ln(\Phi_m/4\pi)$, which is constant for any solution of (5.38). Solutions with $\widetilde{\mathcal{E}} > -3/2$ are dotted, those with $\widetilde{\mathcal{E}} < -3/2$ dashed; both critical solutions are characterised by $\widetilde{\mathcal{E}} = -3/2$. The subsonic dotted solutions represent the 'solar breeze'.

Boundary conditions

$\varrho|_{r_0}$ arbitrary, fixed;

Determine $u|_{r_0}$ from $\Phi_{\mathrm{m},0}{=}\Phi_{\mathrm{m},1}$ or $\left.\dfrac{d\ln\Phi_\mathrm{m}}{dr}\right|_{r_0} {=}0.$

Note that density does not enter Eq. (5.38) — the velocity profile is independent from density. In fact, for the isothermal Parker wind model, if $u(r), \varrho(r)$ is a solution, then $u(r), \alpha\varrho(r)$ is a solution as well, where α is an arbitrary factor. This would be different for other equations of state, e. g. a polytropic solar wind model, where the sound speed depends on ϱ. While such polytropic or isentropic models lead to some qualitative changes (mostly because the critical point is no longer fixed, which introduces an additional degree of freedom), the overall picture (flow starts subsonic, goes through critical point, ends highly supersonic) remains the same.

A flow that shares many similarities with the Parker wind is that of gas through a *Laval nozzle* (see Fig. 5.4 for an illustration), where subsonic gas enters the nozzle, becomes sonic at the point with the smallest cross section (the critical point of the corresponding differential equation) and leaves the nozzle at supersonic speed. In the case of the solar wind, the "nozzle" is provided by the

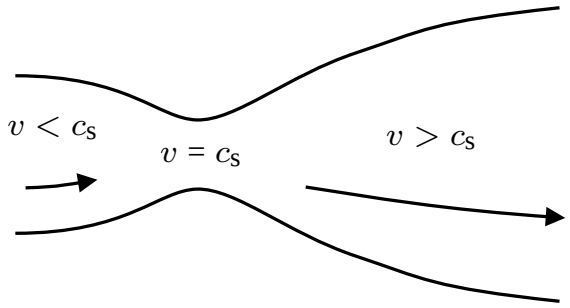

Figure 5.4: Flow through a Laval nozzle (schematic). Subsonic gas from the combustion chamber becomes sonic ($v = c_s$) at the narrowest part of the nozzle and eventually leaves at supersonic speed.

combination of gravity and spherical divergence of streamlines.

Boundary conditions
In a simplified model of the solar wind, one can choose the following boundary conditions at the inner radial boundary $r = r_0$ (the solar surface):

$$\varrho\big|_{r_0} = \varrho_0 \ , \qquad \partial_r \Phi_m\big|_{r_0} = 0 \ , \tag{5.41}$$

where ϱ_0 is given. Written as a condition for u (assuming ϱ to be known), the second condition reads

$$\partial_r v_r = -v_r \left(\partial_r \varrho + \frac{2}{r} \right) \ . \tag{5.42}$$

[Boundary conditions: $\varrho|_{r0}$ arbitrary, fixed; determine $u|_{r_0}$ from $\partial_r \Phi_m|_{r_0} = 0$.]

Chapter 6

Linear and Nonlinear Alfvén Waves

6.1 Basics of magnetohydrodynamics

6.1.1 The induction equation

Magnetohydrodynamics or *MHD* is the discipline describing the *dynamics* of electrically conducting fluids. For such fluids one normally finds that in Ampère's law

$$\operatorname{curl} \mathbf{B} = \mu_0 \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \right) \tag{6.1}$$

the displacement current

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{1}{\mu_0 c^2} \frac{\partial \mathbf{E}}{\partial t}$$

can be neglected. This excludes electromagnetic waves and is equivalent to taking the limit $c \to \infty$. Ampère's law then reduces to

$$\mu_0 \mathbf{j} = \operatorname{curl} \mathbf{B} \ . \tag{6.2}$$

For the permeability we shall take the vacuum value, $\mu_0 = 4\pi \times 10^{-7}$ Vs/Am.

Assuming that in the local rest frame the familiar linear relation $\mathbf{j} = \sigma \mathbf{E}$ between electric current density $\mathbf{j}$ and electric field strength $\mathbf{E}$ holds – σ denoting the electrical conductivity of the medium – and using the Galilei transformation properties of $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{j}$, one arrives at *Ohm's law*

$$\mathbf{j} = \sigma (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \ . \tag{6.3}$$

Thus, like the current the electric field can be expressed in terms of $\mathbf{B}$:

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} + \eta \operatorname{curl} \mathbf{B} \ , \tag{6.4}$$

where

$$\eta \equiv \frac{1}{\mu_0 \sigma} \tag{6.5}$$

is the *magnetic diffusivity* of the medium. Good conductors will have low values of η, while insulators are described by the limit $\eta \to \infty$.

Given that all electromagnetic variables can be explicitly expressed through the magnetic field alone, Faraday's law $\partial \mathbf{B}/\partial t = -\operatorname{curl} \mathbf{E}$ takes the form

$$\frac{\partial \mathbf{B}}{\partial t} = \operatorname{curl} (\mathbf{v} \times \mathbf{B} - \eta \operatorname{curl} \mathbf{B}) \ . \tag{6.6}$$

This equation is referred to as *induction equation*; together with the standard equations of fluid dynamics, it forms the starting point of magnetohydrodynamics.

The term $\operatorname{curl}(\mathbf{v} \times \mathbf{B})$ on the right of (6.6) describes the inducing effect of the motion upon the magnetic field; the term $-\operatorname{curl}(\eta \operatorname{curl} \mathbf{B})$ manifests the "ohmic" decay of the field due to the finite electrical resistance. By order of magnitude we compare the two terms by replacing the diverse vectors by their absolute magnitudes, B and v, and the curl operator by $1/l$, where l is the scale of the field variation in space. We obtain

$$R_{\mathrm{m}} = vl/\eta \tag{6.7}$$

as the ratio of the induction term over the decay term; R_{m} is called the *magnetic Reynolds number*. This number can also be understood as a ratio of two time scales, namely the time scale of ohmic decay, l^2/η, and the advection time scale, l/v. When we speak of "high conductivity" we mean $R_{\mathrm{m}} \gg 1$, or that the time scale of ohmic decay is much longer than l/v.

The magnetic field is always solenoidal,

$$\operatorname{div} \mathbf{B} = 0 \ . \tag{6.8}$$

However, this is not an independent condition: from the induction equation (6.6) it follows that B will remain solenoidal if it is solenoidal at an initial time.

6.1.2 The equations of magnetohydrodynamics

The equations governing the evolution of the magnetic field are supplemented by the equations of fluid dynamics, which have been used already in Sect. 5.1. The continuity equation,

$$\frac{\partial \varrho}{\partial t} = -\operatorname{div}(\varrho \mathbf{v}) \ , \tag{6.9}$$

describes the conservation of matter, and the equation of motion,

$$\varrho \frac{d\mathbf{v}}{dt} = -\operatorname{grad} p + \nu \varrho \left(\Delta \mathbf{v} + \frac{1}{3} \operatorname{grad} \operatorname{div} \mathbf{v} \right) + \varrho \mathbf{g} + \mathbf{j} \times \mathbf{B} \ , \tag{6.10}$$

expresses the conservation of momentum. Equation (5.12) has been multiplied by ϱ; the external accelerations $\mathbf{f}_{\mathrm{ext}}$ represent the gravitational force and the electromagnetic force. The latter – the last term on the right of (6.10) – is the *Lorentz force*. This force (more precisely, force *density*) describes the action of

the magnetic field upon the flow v. The Lorentz force can be written in diverse vectorial forms,

$$\mathbf{j} \times \mathbf{B} = \frac{1}{\mu_0} \operatorname{curl} \mathbf{B} \times \mathbf{B} = -\operatorname{grad} \frac{B^2}{2\mu_0} + (\mathbf{B} \cdot \operatorname{grad}) \frac{\mathbf{B}}{\mu_0} , \tag{6.11}$$

or in form of the divergence of the *Maxwell stress* tensor s_{ik},

$$-\frac{\partial}{\partial x_k} s_{ik} , \qquad s_{ik} = \frac{1}{\mu_0} \left(\frac{1}{2} \delta_{ik} B^2 - B_i B_k \right) . \tag{6.12}$$

The two terms of the Maxwell stress tensor represent the *magnetic pressure*, $B^2/2\mu_0$, and the *magnetic tension*, respectively. The divergence term is a sum over k.

In addition to equations (6.6), (6.9), and (6.10), an energy equation is generally required. However, in the present chapter we shall treat only cases where this equation does not matter, or where it can be represented in a simple way, e.g., in the case of adiabatic changes of state. In that case a fluid parcel does not exchange energy with its environment, so that its pressure and density variations are related through

$$\frac{1}{p} \frac{dp}{dt} = \frac{\gamma}{\varrho} \frac{d\varrho}{dt} , \tag{6.13}$$

where $\gamma = c_p/c_V$. We shall take $\gamma =$ const. and assume that the fluid is an ideal gas. Then

$$\frac{p}{\varrho} = \frac{\mathcal{R}T}{\mu} , \tag{6.14}$$

where μ is the mean molecular weight, and $\mathcal{R}$ the gas constant. The sound speed is defined by

$$c_{\mathrm{s}}^2 := \left(\frac{\partial p}{\partial \varrho} \right)_{\mathrm{S}} = \frac{\gamma \mathcal{R} T}{\mu} = \frac{\gamma p}{\varrho} . \tag{6.15}$$

Expanding the total derivatives in Eq. (6.13) we therefore obtain

$$\frac{\partial p}{\partial t} + \mathbf{v} \cdot \operatorname{grad} p = c_{\mathrm{s}}^2 \left(\frac{\partial \varrho}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \varrho \right) . \tag{6.16}$$

6.1.3 Frozen-in magnetic field

In astrophysics we often deal with ionized, and therefore electrically conducting, matter. Since in most cases the scales are very large, the magnetic Reynolds number R_{m} is large, even when the conductivity is moderate in comparison to the conductors in a terrestrial laboratory. In the limit of large R_{m} the concept of a *frozen magnetic field* is useful.

The meaning of a frozen field is that the magnetic flux is transported along with the material motion. To see this take the total flux

$$\Phi = \int_F \mathbf{B} \cdot \mathbf{df} \tag{6.17}$$

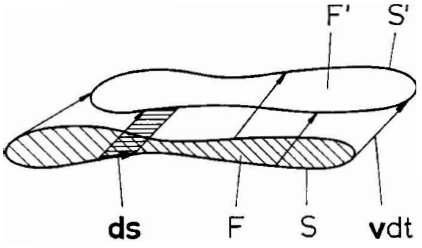

Figure 6.1: Motion of a closed curve S with a flow $\mathbf{v}$.

across an area F, and consider the circumference S of F at two instants separated by an infinitesimal interval dt of time (Fig. 6.1). At the later instant the circumference is made up by the same material "particles" (in the sense of fluid mechanics) but it may be deformed or stretched according to the various paths $\mathbf{v}dt$ followed by these particles. The magnetic flux Φ' across the new area F' may differ from Φ either because the field $\mathbf{B}$ itself has undergone a change in time, or because some of the flux has left the volume between F and F' through the side walls. The (outwards directed) area element of the side wall is $-\mathbf{v}dt \times \mathrm{ds}$; hence

$$\Phi' - \Phi = dt \left[\int_F \dot{\mathbf{B}} \cdot \mathbf{df} + \int_S \mathbf{B} \cdot (\mathbf{v} \times \mathbf{ds}) \right] , \tag{6.18}$$

where the second integral is taken along the closed curve S. Now we use $\mathbf{B} \cdot (\mathbf{v} \times \mathrm{ds}) = (\mathbf{B} \times \mathbf{v}) \cdot \mathrm{ds}$, apply Stokes' theorem to convert the line integral along S into an integral over F, take the limit $dt \to 0$, and obtain

$$\frac{d\Phi}{dt} = \int_F \left[\dot{\mathbf{B}} - \mathrm{curl}(\mathbf{v} \times \mathbf{B}) \right] \cdot \mathbf{df} \ . \tag{6.19}$$

By (6.6), the right-hand side is zero for $\sigma \to \infty$. Hence, in the limit of infinite conductivity, the magnetic flux enclosed by the curve S is conserved. Since the choice of S is arbitrary, and since the constant flux can be represented by a certain number of field lines, we may say that the field lines behave as if they were firmly attached, or "frozen", to the moving fluid.

6.2 Alfvén waves

In a magnetised, electrically conducting fluid, three types of waves can propagate, *slow* and *fast magneto-sonic waves* – which involve compression of the fluid and are related to sound waves –, and *Alfvén waves*. The latter propagate along the magnetic field lines; the restoring force causing them is the tension of the magnetic field and can be expressed in terms of the Maxwell stress tensor s_{ik}.

In this section we shall first discuss the Alfvén waves in a case where even a nonlinear treatment is possible. Then we shall turn to the general case where all three wave types occur; but this will be done only for waves of small amplitude, so that the wave equations can be linearized.

6.2.1 Nonlinear Alfvén waves – no dissipation

This sub-section essentially follows the text of P. H. Roberts, *An Introduction to Magnetohydrodynamics* (Longmans 1967). We concentrate on the Alfvén waves, and neglect the possibility of sound waves. Therefore we neglect the compressibility, and set $\varrho = \text{const.}$. The equation of continuity is, then,

$$\operatorname{div} \mathbf{v} = 0 \; . \tag{6.20}$$

We consider a reference state at rest, $\mathbf{v} = 0$, with a homogeneous magnetic field $\mathbf{B}_0 = \text{const.}$, and write $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$ for the total field strength. The term $\mathbf{b}$ is the variable part of the magnetic field and is associated with the variable flow field $\mathbf{v}$. In the equation of motion, (6.10), we neglect the forces due to viscosity and gravity, expand the derivative $d\mathbf{v}/dt$, substitute $\mathbf{B}$, and use (6.11):

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \mathbf{v} = -\frac{1}{\varrho} \operatorname{grad} \left(p + \frac{B^2}{2\mu_0} \right) + \frac{1}{\mu_0 \varrho} \mathbf{B}_0 \cdot \operatorname{grad} \mathbf{b} + \frac{1}{\mu_0 \varrho} \mathbf{b} \cdot \operatorname{grad} \mathbf{b} \; . \tag{6.21}$$

In the induction equation, (6.6), we neglect the η term (which corresponds to the limit $R_\mathrm{m} \gg 1$), expand the term $\operatorname{curl}(\mathbf{v} \times \mathbf{B})$, and also substitute $\mathbf{B}$:

$$\frac{\partial \mathbf{b}}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \mathbf{b} = \mathbf{B}_0 \cdot \operatorname{grad} \mathbf{v} + \mathbf{b} \cdot \operatorname{grad} \mathbf{v} \; . \tag{6.22}$$

Now suppose there is a solution satisfying $p + B^2/2\mu_0 = \text{const.}$ (this condition can be used to determine the pressure p from B^2, once a solution for the magnetic field has been found). In such a case both equations (6.21) and (6.22) can be satisfied by

$$\mathbf{v} = \pm \frac{\mathbf{b}}{\sqrt{\mu_0 \varrho}} \; , \tag{6.23}$$

because they both reduce to

$$\frac{\partial \mathbf{b}}{\partial t} = \pm \frac{\mathbf{B}_0}{\sqrt{\mu_0 \varrho}} \cdot \operatorname{grad} \mathbf{b} \; . \tag{6.24}$$

We choose a Cartesian coordinate system such that $\mathbf{B}_0$ is in the x-direction, and define the *Alfvén velocity*:

$$v_\mathrm{A} = \frac{B_0}{\sqrt{\mu_0 \varrho}} \; . \tag{6.25}$$

Then, (6.24) becomes

$$\frac{\partial \mathbf{b}}{\partial t} = \pm v_\mathrm{A} \frac{\partial \mathbf{b}}{\partial x} \tag{6.26}$$

and has solutions of the form

$$\mathbf{b} = B_0 \mathbf{f}(x \pm v_\mathrm{A} t, y, z) \; . \tag{6.27}$$

The factor B_0 has been added to make $\mathbf{f}$ a dimensionless function, which otherwise is quite arbitrary except that, because $\mathbf{b}$ is solenoidal, it must satisfy $\operatorname{div} \mathbf{f} = 0$. From (6.23) it follows that

$$\mathbf{v} = \pm v_\mathrm{A} \mathbf{f}(x \pm v_\mathrm{A} t, y, z) \; . \tag{6.28}$$

Remark 1: Alfvén-wave energy
The solution (6.27) and (6.28) propagates either in the negative x-direction (upper sign) or in the positive x-direction (lower sign). These solutions are called *Alfvén waves* (Hannes Alfvén, Nobel price in physics 1975). Notice that the two solutions cannot be superposed into a linear combination. Such a superposition would *not* solve the (nonlinear!) equations (6.21) and (6.22).

Let us consider the solution that propagates in the positive x-direction. And let us consider the special simple case where $\mathbf{b}$ and $\mathbf{v}$ are independent of y and z. From $\operatorname{div}\mathbf{f} = 0$ we then find $\partial f_x/\partial x = 0$. Hence f_x must be constant; we set $f_x = 0$ as a non-zero value could be added to the equilibrium field B_0. Thus $b_x = v_x = 0$, i. e., the propagating pattern is *transverse*.

The energy density (magnetic plus kinetic) is

$$\frac{1}{2\mu_0}(\mathbf{B}_0 + \mathbf{b})^2 + \frac{1}{2}\varrho v^2 = \frac{1}{2\mu_0}(B_0^2 + b^2) + \frac{1}{2}\varrho v^2 \ , \tag{6.29}$$

and the part associated with the Alfvén wave is

$$e_{\text{wave}} = \frac{1}{2\mu_0}b^2 + \frac{1}{2}\varrho v^2 = \varrho v_{\text{A}}^2 f^2 \ . \tag{6.30}$$

The energy flux associated with the Alfvén wave can be obtained from the Poynting vector. Since, by (6.4), the electric field is $\mathbf{E} = -\mathbf{v}\times\mathbf{B} = v_{\text{A}}\mathbf{f}\times\mathbf{B}_0$, the Poynting vector is

$$\mathbf{P} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = \frac{v_{\text{A}}}{\mu_0}(\mathbf{f}\times\mathbf{B}_0)\times(\mathbf{B}_0 + \mathbf{b}) \ . \tag{6.31}$$

Expanding the double vector product and taking into account that $\mathbf{b}$ is transverse, we find

$$\mathbf{P} = \varrho v_{\text{A}}^2\mathbf{v} + v_{\text{A}}e_{\text{wave}}\mathbf{e}_x \ . \tag{6.32}$$

For a periodic oscillation $\mathbf{v}$ the first term makes no contribution to the average over a period. On the other hand, the second term represents a net transport in the positive x-direction. This is called *Alfvén radiation*. This kind of energy transport probably plays an essential part in the heating of the solar corona, where energy must be transmitted from a cool (the solar surface, approx. 5800 K) to a hot place (the corona, over 10^6 K), which usually does not happen in thermodynamics.

Remark 2: The string analogy
Let y be the direction of oscillation, i.e., $\mathbf{v} = (0,\, v_y(x),\, 0)$ and $\mathbf{b} = (0,\, b_y(x),\, 0)$. The y-component of (6.21) then reads

$$\varrho\frac{\partial v_y}{\partial t} = \frac{1}{\mu_0}B_0\frac{\partial b_y}{\partial x} \ . \tag{6.33}$$

The force on the right can be written as

$$\frac{1}{\mu_0}\frac{\partial B_0 b_y}{\partial x} = \frac{1}{\mu_0}\frac{\partial B_x B_y}{\partial x} = -\frac{\partial s_{xy}}{\partial x} \ , \tag{6.34}$$

cf. Eq. (6.12). On the other hand, the restoring force of an oscillating string is given by $\partial(T\sin\alpha)/\partial x$, where $T\sin\alpha$ is the transverse component of the string tension (the string has an angle $\alpha(x,t)$ to the x-direction). This analogy justifies the name "magnetic tension" of the component s_{xy} of the Maxwell stress tensor.

The time derivative of (6.24) is

$$\frac{\partial^2 b_y}{\partial t^2} = \pm\frac{B_0}{\sqrt{\mu_0\varrho}}\frac{\partial}{\partial x}\frac{\partial b_y}{\partial t} = \frac{B_0^2}{\mu_0\varrho}\frac{\partial^2 b_y}{\partial x^2} = v_{\rm A}^2\frac{\partial^2 b_y}{\partial x^2}\ . \tag{6.35}$$

Hence $v_{\rm A}$ is the propagation speed of the transverse field oscillation, corresponding to $(T/\varrho)^{1/2}$ for the oscillating string.

6.2.2 Alfvén waves with dissipation

In this section we retain the assumption of incompressibility and transverse wave motion, but we admit dissipation, that is, we admit the η and ν terms in Eqs. (6.6) and (6.10), respectively. With the same choice of coordinates as in the preceding section, we find for $v_y(x,t)$ and $b_y(x,t)$ the evolution equations

$$\varrho\frac{\partial v_y}{\partial t} = \frac{B_0}{\mu_0}\frac{\partial b_y}{\partial x} + \nu\varrho\frac{\partial^2 v_y}{\partial x^2}\ , \tag{6.36}$$

$$\frac{\partial b_y}{\partial t} = B_0\frac{\partial v_y}{\partial x} + \eta\frac{\partial^2 b_y}{\partial x^2}\ . \tag{6.37}$$

Roberts (*l.c.*) considers a wave excited by a harmonic oscillation $\exp(i\omega t)$ at the boundary $x=0$ and inquires how far in space x the damped oscillation reaches. He finds $L_{\rm D} = \lambda^2 v_{\rm A}/[2\pi^2(\nu+\eta)]$ for the length of amplitude decay (to 1/e of the value at $x=0$), where $\lambda = 2\pi v_{\rm A}/\omega$ is the wave length of the propagating Alfvén wave. Here we proceed in a different, but equivalent, way. We consider perturbations $v_y,\ b_y \propto \exp[i(kx-\omega t)]$, where k is a given real wave number, while ω must be found (and will be complex in general). Substitution into (6.36) and (6.37) yields two homogeneous equations for the amplitudes of v_y and b_y. For a non-trivial solution the determinant of this system of equations must vanish. This yields an equation for ω, namely

$$\omega^2 + i\omega k^2(\nu+\eta) - \nu\eta k^4 - v_{\rm A}^2 k^2 = 0\ . \tag{6.38}$$

The roots of this equation are

$$\omega_\pm = \frac{1}{2}\left(-ik^2(\nu+\eta) \pm \sqrt{4v_{\rm A}^2k^2 - k^4(\nu-\eta)^2}\right)\ . \tag{6.39}$$

Now we consider the case of *weak damping*, with $\nu,\eta \ll v_{\rm A}/k$. Then

$$\omega_\pm = \pm k v_{\rm A} - ik^2\frac{\nu+\eta}{2}\ , \tag{6.40}$$

and the perturbation is a damped Alfvén wave, with the (1/e)-time

$$T_{\rm D} = -\frac{1}{\Im(\omega)} = \frac{2}{k^2(\nu+\eta)} = \frac{\lambda^2}{2\pi^2(\nu+\eta)} \ , \tag{6.41}$$

where $\lambda = 2\pi/k$. Within the time $T_{\rm D}$ the wave propagates a distance $v_{\rm A}T_{\rm D}$, i.e., the *damping length* is

$$L_{\rm D} = \frac{v_{\rm A}\lambda^2}{2\pi^2(\nu+\eta)} \ . \tag{6.42}$$

6.2.3 Linear Alfvén waves and magneto-sonic waves

Now we retain compressibility, and allow for wave propagation in an arbitrary direction (relative to the direction of B_0). But we shall discard the dissipation terms (the ν and η terms). And we shall consider only waves of small amplitude, relative to the reference state at rest, $\mathbf{v} = 0$:

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{b} \ , \quad \varrho = \varrho_0 + \varrho_1 \ , \quad p = p_0 + p_1 \ . \tag{6.43}$$

We consider an isothermal situation, for the reference state as well as for the perturbation, and therefore have $p_1 = c_{\rm s}^2\varrho_1$. Since B_0 is homogeneous, the associated current density is zero, and

$$\mathbf{j} = \mathbf{j}_1 = \frac{1}{\mu_0}\,\mathrm{curl}\,\mathbf{b} \ . \tag{6.44}$$

Substituting all variables in Eqs. (6.6), (6.9), and (6.10), and retaining only the terms of first order in $\mathbf{v}$, $\mathbf{b}$, etc., we find

$$\frac{\partial\varrho_1}{\partial t} = -\varrho_0\,\mathrm{div}\,\mathbf{v} \ , \tag{6.45}$$

$$\varrho_0\frac{\partial\mathbf{v}}{\partial t} = -c_{\rm s}^2\mathrm{grad}\,\varrho_1 + \frac{1}{\mu_0}\,\mathrm{curl}\,\mathbf{b}\times\mathbf{B}_0 \ , \tag{6.46}$$

$$\frac{\partial\mathbf{b}}{\partial t} = \mathrm{curl}\,(\mathbf{v}\times\mathbf{B}_0) \ . \tag{6.47}$$

Since all coefficients in these linear equations are constants, we seek solutions $\propto \exp[i(\mathbf{k}\cdot\mathbf{x} - \omega t)]$. Hence we may substitute

$$\frac{\partial}{\partial t} = -i\omega \ , \quad \mathrm{grad} = i\mathbf{k} \ , \quad \mathrm{div} = i\mathbf{k}\cdot \ , \quad \mathrm{curl} = i\mathbf{k}\times \tag{6.48}$$

and obtain

$$-i\omega\varrho_1 = -\varrho_0 i\mathbf{k}\cdot\mathbf{v},, \tag{6.49}$$

$$-\varrho_0 i\omega\mathbf{v} = -c_{\rm s}^2 i\mathbf{k}\varrho_1 + \frac{1}{\mu_0}i(\mathbf{k}\times\mathbf{b})\times\mathbf{B}_0 \ , \tag{6.50}$$

$$-i\omega\mathbf{b} = i\mathbf{k}\times(\mathbf{v}\times\mathbf{B}_0) \ . \tag{6.51}$$

We eliminate ϱ_1 with the first, and $\mathfrak{b}$ with the third of these equations. The second, multiplied by $i\omega/\varrho_0$, then yields

$$\begin{aligned}\omega^2\mathbf{v} &= c_\mathrm{s}^2\mathbf{k}\,(\mathbf{k}\cdot\mathbf{v}) + \frac{1}{\varrho_0\mu_0}\mathbf{B}_0\times\{\mathbf{k}\times[\mathbf{k}\times(\mathbf{B}_0\times\mathbf{v})]\} \\ &= c_\mathrm{s}^2\mathbf{k}\,(\mathbf{k}\cdot\mathbf{v}) + \frac{1}{\varrho_0\mu_0}\{\mathbf{k}\,[B_0^2\,(\mathbf{k}\cdot\mathbf{v}) - (\mathbf{v}\cdot\mathbf{B}_0)(\mathbf{k}\cdot\mathbf{B}_0)] \\ &\qquad -(\mathbf{k}\cdot\mathbf{B}_0)\,[\mathbf{B}_0\,(\mathbf{k}\cdot\mathbf{v}) - \mathbf{v}\,(\mathbf{k}\cdot\mathbf{B}_0)]\}\,. \end{aligned} \tag{6.52}$$

We take the scalar product of this equation with the vectors $\mathbf{k}$ and $\mathbf{B}_0$ (i.e., the projection of the equation in these two directions),

$$\omega^2\,(\mathbf{k}\cdot\mathbf{v}) = c_\mathrm{s}^2k^2\,(\mathbf{k}\cdot\mathbf{v}) + \frac{1}{\varrho_0\mu_0}\{k^2[B_0^2\,(\mathbf{k}\cdot\mathbf{v}) - (\mathbf{v}\cdot\mathbf{B}_0)(\mathbf{k}\cdot\mathbf{B}_0)]\}\,, \tag{6.53}$$

$$\omega^2\,(\mathbf{B}_0\cdot\mathbf{v}) = c_\mathrm{s}^2(\mathbf{k}\cdot\mathbf{B}_0)(\mathbf{k}\cdot\mathbf{v})\,. \tag{6.54}$$

These are two homogeneous linear equations for the two unknowns $\mathbf{k}\cdot\mathbf{v}$ and $\mathbf{B}_0\cdot\mathbf{v}$; their determinant must be zero. This yields the *dispersion relation* for the magneto-sonic waves:

$$\omega^4 - \omega^2\,(c_\mathrm{s}^2 + v_\mathrm{A}^2)k^2 + \frac{c_\mathrm{s}^2k^2}{\varrho_0\mu_0}(\mathbf{k}\cdot\mathbf{B}_0)^2 = 0. \tag{6.55}$$

Let θ be the angle between the direction $\mathbf{k}$ of wave propagation and $\mathbf{B}_0$. Then $(\mathbf{k}\cdot\mathbf{B}_0)^2/(\varrho_0\mu_0) = k^2v_\mathrm{A}^2\cos^2\theta$, and the dispersion relation becomes

$$\frac{\omega^4}{k^4} - (c_\mathrm{s}^2 + v_\mathrm{A}^2)\frac{\omega^2}{k^2} + c_\mathrm{s}^2v_\mathrm{A}^2\cos^2\theta = 0\,. \tag{6.56}$$

The solution is

$$\frac{\omega^2}{k^2} = \frac{c_\mathrm{s}^2 + v_\mathrm{A}^2 \pm \sqrt{(c_\mathrm{s}^2 + v_\mathrm{A}^2)^2 - 4c_\mathrm{s}^2v_\mathrm{A}^2\cos^2\theta}}{2}\,, \tag{6.57}$$

and the phase velocity, ω/k, is the square root of this.

The upper sign in (6.57) yields the *fast magneto-sonic wave*; its phase velocity is largest, namely $\omega^2/k^2 = c_\mathrm{s}^2 + v_\mathrm{A}^2$, for $\theta = 90^\circ$, i.e., for a wave propagating perpendicular to the reference field $\mathbf{B}_0$; it is slowest, $\omega^2/k^2 = \max(c_\mathrm{s}^2, v_\mathrm{A}^2)$, for $\theta = 0$.

The lower sign in (6.57) yields the *slow magneto-sonic wave*; its phase velocity is smallest for $\theta = 90^\circ$, namely $\omega/k = 0$; it reaches its largest value for $\theta = 0$, where $\omega^2/k^2 = \min(c_\mathrm{s}^2, v_\mathrm{A}^2)$.

Figure 6.2 is a polar diagram that shows the phase velocities of the two magneto-sonic waves, as functions of the angle θ, for the cases $v_\mathrm{A}^2 > c_\mathrm{s}^2$ and $v_\mathrm{A}^2 < c_\mathrm{s}^2$.

Above we have projected Eq. (6.52) onto the two directions $\mathbf{k}$ and $\mathbf{B}_0$. A third independent direction is $\mathbf{k}\times\mathbf{B}_0$. Projection of (6.52) to this direction yields

$$\omega^2\,(\mathbf{k}\times\mathbf{B}_0)\cdot\mathbf{v} = \frac{1}{\varrho_0\mu_0}(\mathbf{k}\cdot\mathbf{B}_0)^2\,(\mathbf{k}\times\mathbf{B}_0)\cdot\mathbf{v} \tag{6.58}$$

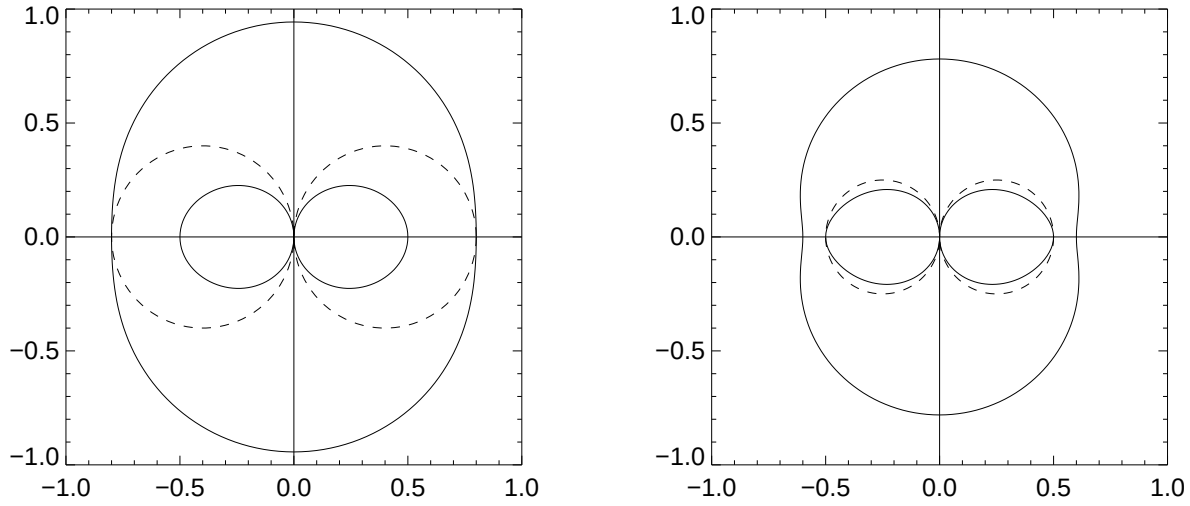

Figure 6.2: Polar diagram for the dispersion of magneto-sonic waves (*solid*) and Alfvén waves (*dashed*). The field B_0 points toward the right. The distance from the origin is the phase velocity, ω/k, as a function of the angle θ between the field and the direction of wave propagation. *Left*: $v_\mathrm{A} = 0.8$, $c_\mathrm{s} = 0.5$. *Right*: $v_\mathrm{A} = 0.5$, $c_\mathrm{s} = 0.6$.

or

$$\frac{\omega^2}{k^2} = v_\mathrm{A}^2 \cos^2\theta \; . \tag{6.59}$$

This is the dispersion relation for an Alfvén wave that propagates in an inclined direction relative to the reference field B_0; for $\theta = 0$ it includes the special case treated earlier. Relation (6.59) is also illustrated in Fig. 6.2.

In this section we have found two independent wave types: On the one hand, the fast and slow magneto-sonic waves, with velocity components in the plane defined by the vectors k and B_0; on the other hand, the inclined Alfvén wave, with a velocity component perpendicular to both k and B_0. Although the latter does not, in general, propagate along the field B_0, it is a strictly transverse wave.

Chapter 7

Wave Breaking and Shocks

7.1 Sound waves in a stratified atmosphere

In this section we shall demonstrate that in a stratified atmosphere acoustic waves have an amplitude that increases with height. Hence the formation of shocks is inevitable. Such shocks will be discussed in Sect. 7.2.

Stratification is a consequence of gravity. Therefore we retain the term $\varrho \mathbf{g} = -\varrho g \mathbf{e}_z$ in Eq. (6.10). On the other hand, we neglect the effects of viscosity and of the magnetic field, and consider only waves of small amplitude, and adiabatic changes of state, as expressed by Eq. (6.16).

We consider an isothermal equilibrium state at rest, $\mathbf{v} = 0$. In this state the pressure and the density have *barometric* stratifications,

$$p_0(z) = p_{00} e^{-z/H} , \qquad \varrho(z) = \varrho_{00} e^{-z/H} , \tag{7.1}$$

where the *scale height*

$$H = \frac{\mathcal{R}T}{\mu g} = \frac{c_{\mathrm{s}}^2}{\gamma g} \tag{7.2}$$

is a constant. Small deviations $\mathbf{v}$, p_1, ϱ_1 from the equilibrium state obey the linearized forms of Eqs. (6.9), (6.10), and (6.16):

$$\frac{\partial \varrho_1}{\partial t} = -\varrho_0 \operatorname{div} \mathbf{v} - \mathbf{v} \cdot \operatorname{grad} \varrho_0 , \tag{7.3}$$

$$\varrho_0 \frac{\partial \mathbf{v}}{\partial t} = -\operatorname{grad} p_1 - \varrho_1 g \mathbf{e}_z , \tag{7.4}$$

$$\frac{\partial p_1}{\partial t} + v_z \frac{dp_0}{dz} = c_{\mathrm{s}}^2 \left(\frac{\partial \varrho_1}{\partial t} + v_z \frac{d\varrho_0}{dz} \right) . \tag{7.5}$$

These equations are linear and homogeneous, but this time the coefficients are *not* constants: ϱ_0 and p_0 depend on z. A close inspection, however, shows that the z-dependence has always the form $\exp(-z/H)$, and that it occurs always in combination with the velocity. Therefore the ansatz

$$(p_1,\, \varrho_1) = (p_{10},\, \varrho_{10})\, e^{-z/2H} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)] , \tag{7.6}$$

$$\mathbf{v} = \mathbf{v}_0\, e^{z/2H} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \tag{7.7}$$

yields a factor $\exp(-z/2H)$ in *all* terms of the three equations; hence this factor can be dropped, as well as the factor $\exp[i(\mathbf{k}\cdot\mathbf{x}-\omega t)]$. Substitution of the ansatz leads to a linear homogeneous algebraic system of equations for the amplitudes p_{10}, ϱ_{10}, and $\mathbf{v}_0$, viz.

$$-i\omega\varrho_{10} = -\varrho_{00}\left(i\mathbf{k}+\frac{\mathbf{e}_z}{2H}\right)\cdot\mathbf{v}_0 + v_{z0}\frac{\varrho_{00}}{H} , \tag{7.8}$$

$$-i\omega\varrho_{00}\mathbf{v}_0 = -\left(i\mathbf{k}-\frac{\mathbf{e}_z}{2H}\right)p_{10} - \varrho_{10}g\mathbf{e}_z , \tag{7.9}$$

$$-i\omega p_{10} - v_{z0}\frac{p_{00}}{H} = c_\mathrm{s}^2\left(-i\omega\varrho_{10} - v_{z0}\frac{\varrho_{00}}{H}\right) . \tag{7.10}$$

It is sufficiently general to treat the two-dimensional case where $\mathbf{v}_0 = (v_{x0},\, 0,\, v_{z0})$ and $\mathbf{k} = (k_x,\, 0,\, k_z)$. Thus we have four homogeneous linear equations for the four unknowns ϱ_{10}, v_{x0}, v_{z0}, and p_{10}; the determinant of these four equations must vanish. We omit the calculation; the result is a fourth-order equation for the frequency ω, the *dispersion relation* for waves in an isothermally stratified atmosphere:

$$\omega^4 - c_\mathrm{s}^2\left(k_x^2+k_z^2+\frac{1}{4H^2}\right)\omega^2 + k_x^2c_\mathrm{s}^2\frac{g}{H}\frac{\gamma-1}{\gamma} = 0 . \tag{7.11}$$

Alternatively, we may consider ω as given and solve for the vertical component, k_z, of the wave vector. Introducing the *acoustic cutoff frequency*

$$\omega_\mathrm{A} = \frac{c_\mathrm{s}}{2H} , \tag{7.12}$$

and the *Brunt–Väisälä frequency* N through

$$N^2 = \frac{g}{H}\frac{\gamma-1}{\gamma} , \tag{7.13}$$

we obtain

$$k_z^2 = \frac{\omega^2-\omega_\mathrm{A}^2}{c_\mathrm{s}^2} + k_x^2\frac{N^2-\omega^2}{\omega^2} . \tag{7.14}$$

Three topics must be mentioned here:

- The dispersion relation is of second order in ω^2. Hence it essentially describes two wave types. One is an acoustic wave, modified by the effect of stratification. The other is an *internal gravity wave*, related to gravity waves in water, and essentially supported by the restoring force of gravity. The internal gravity wave exists only in the case of non-zero horizontal wave number, $k_x \neq 0$.

- The modified acoustic wave is easiest discussed in the special case $k_x = 0$, where

$$k_z^2 = \frac{\omega^2 - \omega_{\rm A}^2}{c_{\rm s}^2} \ . \tag{7.15}$$

  This means that a vertically propagating acoustic wave (i.e., one with real vertical wave number k_z) is possible only if $\omega^2 > \omega_{\rm A}^2$. Otherwise the vertical dependence of the amplitude is exponential. In the solar atmosphere, we have $c_{\rm s} \approx 7\,$km/s and $H \approx 100\,$km; hence $\omega_{\rm A} = c_{\rm s}/2H \approx 0.035\,{\rm s}^{-1}$, which corresponds to a period of about 3 min. Indeed, it appears that only waves with periods shorter than this propagate in the solar atmosphere (Fig. 7.1).

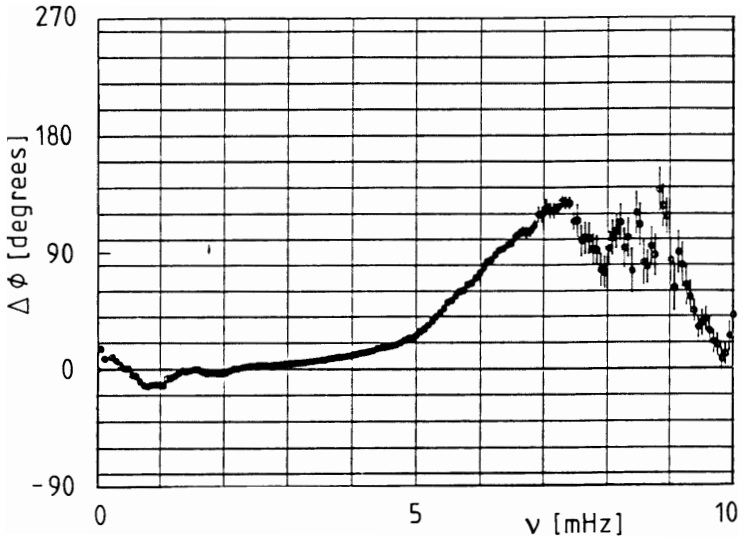

Figure 7.1: The phase difference, $\Delta\phi$, of oscillations measured in two lines originating deep and high, respectively, in the solar atmosphere, as a function of frequency. A significant phase difference, indicating *running waves*, occurs only for frequencies $\nu = \omega/2\pi > 5\,$mHz. After J. Staiger, *Astron. Astrophys.* **175** (1987), p. 263.

- Wave amplitudes increase exponentially with height, proportionally to $\exp(z/2H)$. A typical value for a wave amplitude at the base of the solar atmosphere (the *photosphere*) is 1 km/s, as observed by means of the Doppler effect of spectral lines. With the numbers given before, the sound speed will be exceeded already after 4 scale heights. Then the nonlinear terms that were dropped in the present section become important, and shock fronts will form.

Shocks also occur in other astrophysical contexts. An important example is the solar wind which was treated in Sect. 5.2. The solution derived by Parker is super-sonic after passing the critical point. Thus, whenever this super-sonic wind meets an obstacle like a planet or a comet, there must be a transition to a sub-sonic regime, which will have the form of a shock front. In addition, the solar wind finally must merge into the interstellar gas, which should happen after passing a shock front. According to calculations this *termination shock* of the solar wind occurs at a distance of order 85 astronomical units from the Sun, more than twice the distance of Pluto, the outermost planet. Two spacecrafts, *Voyager 1* and *Voyager 2*, launched over 20 years ago and still able to transmit signals, will reach that distance in 2004 and 2010, respectively.

7.2 The shock-tube problem

7.2.1 The experiment

The shock-tube experiment is performed in a tube filled with gas that is initially at rest. A membrane separates two parts of the tube (Fig. 7.2). The pressure and the density on the left and right of the membrane are p_l, ϱ_l and p_r, ϱ_r, respectively.

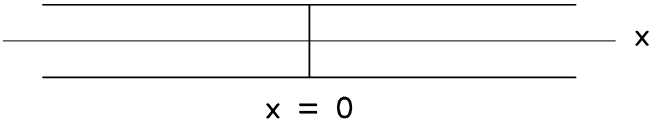

Figure 7.2: The shock-tube, with the membrane at $x = 0$.

To start the experiment, the pressure at the left is increased until the membrane breaks. At this moment a flow towards the right sets in. If the original pressure difference across the membrane is sufficiently large, one observes a forward-moving *shock front*, where the velocity as well as the thermodynamic variables are discontinuous (at $x \approx 1.25$ in Fig. 7.4 below). Energy is dissipated in such a shock, which results in heating of the gas. The shock front is followed by an – also forward-moving – *contact discontinuity* (at $x \approx 0.3$ in Fig. 7.4), where the pressure and the velocity are continuous, while the other thermodynamic variables have jumps. Finally, there is a – backward-moving – *zone of expansion*, also called *rarefaction wave* ($-1.1 < x < -0.6$), with an approximately linear transition to the original high-pressure state at the left side. Figure 7.3 shows the positions of the diverse transitions in the shock-tube experiment as functions of time.

7.2.2 Numerical model

We describe the shock-tube experiment in a one-dimensional model. The coordinate x is along the tube, and the membrane that separates the initial high- and low-pressure regimes is placed at $x = 0$.

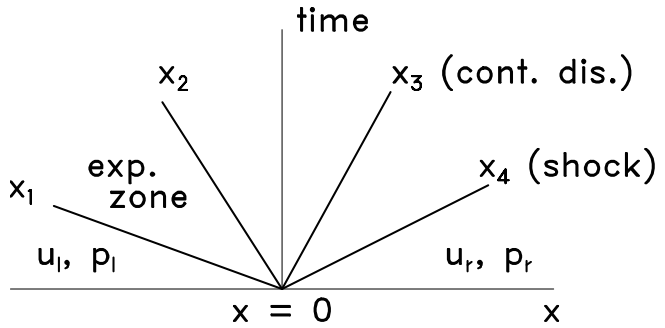

Figure 7.3: The change of position of the diverse transitions in the shock-tube experiment. Exp. zone = expansion zone; cont. dis. = contact discontinuity.

The three equations to be used are the equation of continuity, the equation of motion, and the equation of energy conservation. The variables for which we write down these differential equations are $\ln \varrho$, the velocity v, and the specific entropy s. Other thermodynamic variables occur as well, but merely as auxiliary quantities.

The continuity equation is, cf. (5.19),

$$\frac{\partial \ln \varrho}{\partial t} = -v \frac{\partial \ln \varrho}{\partial x} - \frac{\partial v}{\partial x} \; . \tag{7.16}$$

For the equation of motion we take the form (6.10), but we neglect the terms arising from gravity and the magnetic field. On the other hand, we must recall that in the case of a variable viscosity the coefficient $\nu\varrho$ appears under the divergence of the stress tensor, cf. Eqs. (5.7) and (5.10). In our 1D-model, then, we have

$$\frac{\partial v}{\partial t} = -v \frac{\partial v}{\partial x} - \frac{1}{\varrho} \frac{\partial p}{\partial x} + \frac{1}{\varrho} \frac{\partial}{\partial x} \left(\nu \varrho \frac{4}{3} \frac{\partial v}{\partial x} \right) \; . \tag{7.17}$$

In order to eliminate the pressure from this equation, we start from the definition of the specific entropy (per mass), $s(\epsilon, V)$, as a function of the specific internal energy ϵ and the specific volume $V = 1/\varrho$. Since, for a perfect gas, $\epsilon = c_v T$, we have

$$ds = \frac{1}{T} d\epsilon + \frac{p}{T} dV = \frac{c_v}{T} dT - \frac{p}{T\varrho^2} d\varrho \; . \tag{7.18}$$

Using $p = \varrho \mathcal{R} T/\mu = \varrho T(c_p - c_v)$ and $dT/T = dp/p - d\varrho/\varrho$ we find the desired relation between entropy, pressure, and density:

$$ds = c_v (d \ln p - \gamma d \ln \varrho) \; , \tag{7.19}$$

where $\gamma = c_p/c_v$, or, in integrated form,

$$s = s_0 + c_v \ln \left[\frac{p}{p_0} \left(\frac{\varrho}{\varrho_0} \right)^{-\gamma} \right] \; , \qquad p = p_0 \left(\frac{\varrho}{\varrho_0} \right)^{\gamma} \exp[(s - s_0)/c_v] \; , \tag{7.20}$$

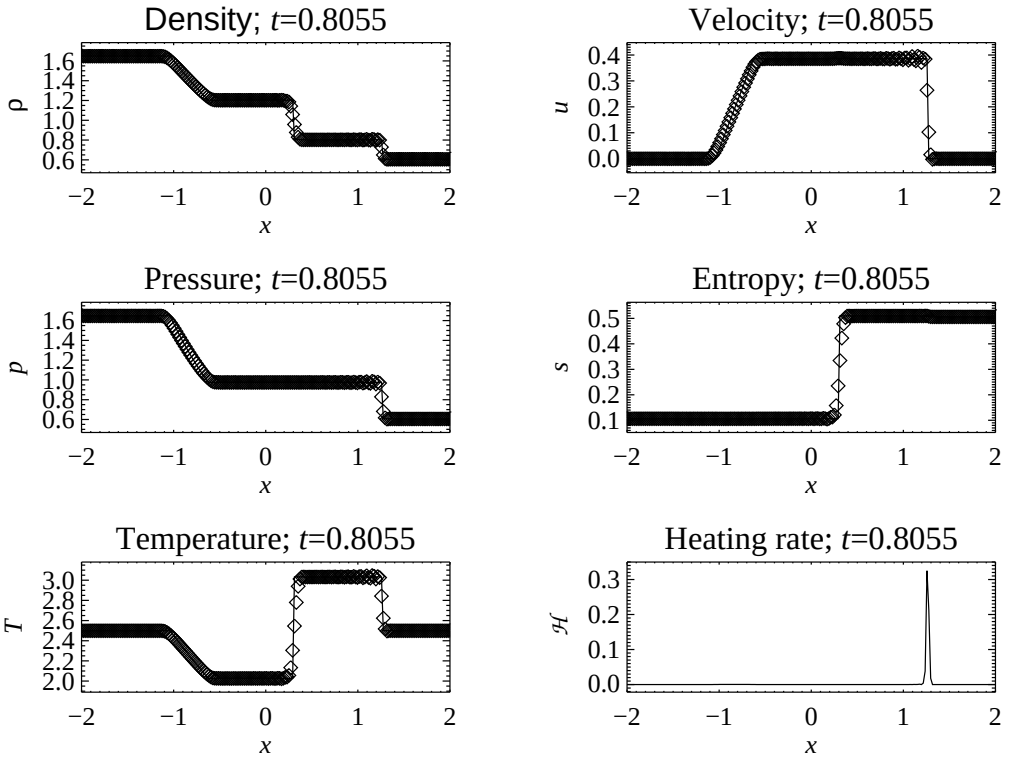

Figure 7.4: The variables of the shock-tube experiment. From the numerical model of Sect. 7.2.2.

where the subscript 0 denotes a reference state. For the sound velocity we thus obtain

$$c_{\mathrm{s}}^2 \equiv \frac{p\gamma}{\varrho} = c_{\mathrm{s}\,0}^2 \left(\frac{\varrho}{\varrho_0}\right)^{\gamma-1} \exp[(s-s_0)/c_v] \; . \tag{7.21}$$

Hence the equation of motion to be used in the numerical model can be written in the form

$$\frac{\partial v}{\partial t} = -v\frac{\partial v}{\partial x} - c_{\mathrm{s}}^2 \frac{\partial}{\partial x}\left(\ln \varrho + \frac{s}{c_p}\right) + \frac{4\nu}{3}\left[\frac{\partial^2 v}{\partial x^2} + \frac{\partial}{\partial x}\left(\ln(\nu\varrho)\frac{\partial v}{\partial x}\right)\right] \; . \tag{7.22}$$

Next we turn to the energy equation.

We begin with the 1st law of thermodynamics, $\varrho T ds = \delta Q$, where δQ is the energy gain (per volume, because a factor ϱ was added on the left). In the shock-tube model, we consider two contributions to δQ, namely heat conduction and dissipation due to viscosity. The flow of conducted heat is down the temperature gradient, $-\lambda\partial/T\partial x$, and its negative divergence is $-\partial/\partial x$ of this. The heat added by dissipation is $\varrho\nu\mathsf{S}^2/2$, where S is the stress tensor, defined through $\mathsf{S}_{ik} = (\partial_i v_k + \partial_k v_i) - (2/3)\delta_{ik}\,\mathrm{div}\,\mathbf{v}$; up to the factor $\varrho\nu$, this is the tensor σ_{ik} introduced in Sect. 5.1. In the one-dimensional model the dissipation term simplifies, so that

$$\varrho T\frac{ds}{dt} = \frac{\partial}{\partial x}\left(\lambda\frac{\partial T}{\partial x}\right) + \frac{4}{3}\varrho\nu\left(\frac{\partial v}{\partial x}\right)^2 \; . \tag{7.23}$$

In order to rewrite the heat-conduction term we replace the heat conductivity λ by the thermal diffusion coefficient

$$\chi = \frac{\lambda}{\varrho c_p} \,. \tag{7.24}$$

The energy equation, divided by ϱT, thus takes the form

$$\frac{\partial s}{\partial t} = -v\frac{\partial s}{\partial x} + c_p\chi\left[\frac{\partial^2 \ln T}{\partial x^2} + \frac{\partial \ln T}{\partial x}\frac{\partial \ln(T\varrho\chi)}{\partial x}\right] + \frac{4\nu}{3T}\left(\frac{\partial v}{\partial x}\right)^2 \,, \tag{7.25}$$

where $d/dt = \partial/\partial t + v\partial/\partial x$ has been used.

Equations (7.16), (7.22), and (7.25) are used in the exercises, part 8, of this course.

7.2.3 Rankine–Hugoniot relations

In this section we consider the diverse discontinuities at the shock front, and derive relations between the physical quantities at the two sides of the shock. We do this for the idealized case without dissipation, where $\nu = \chi = 0$, which is also called the *Riemann problem*. The position of the shock front is at $x_4(t)$, see Fig. 7.3.

We integrate the equation of continuity,

$$\frac{\partial \varrho}{\partial t} + \frac{\partial(\varrho v)}{\partial x} = 0 \,, \tag{7.26}$$

over a short interval that includes x_4 as an inner point, and then go to the limit where the length of that interval becomes zero. Since $\partial\varrho/\partial t$ is bounded, the contribution from this term vanishes in the limit. Since the remaining term is an x-derivative, we can integrate and obtain

$$[[\varrho v]] = 0 \,, \tag{7.27}$$

where $[[\ \ ]]$ denotes the difference of the enclosed expression on the two sides of the discontinuity (right $-$ left). It is important to realize that $x_4(t)$ is moving; hence the velocity that appears in (7.27) is *relative to the (moving) shock front*. Condition (7.27) expresses that as much matter as enters the front on one side will leave the front at the other side.

In the equation of motion we set $\nu = 0$. If we add v times the equation of continuity, we obtain

$$\frac{\partial \varrho v}{\partial t} = -\frac{\partial \varrho v^2}{\partial x} - \frac{\partial p}{\partial x} \,. \tag{7.28}$$

From this equation we find, by the same reasoning as before,

$$[[p + \varrho v^2]] = 0 \,. \tag{7.29}$$

Again the velocity that appears in this relation is relative to the moving shock front.

A third relation is obtained from the energy equation,

$$\frac{\partial}{\partial t}\left(\varrho\epsilon+\frac{1}{2}\varrho v^2\right)=-\frac{\partial}{\partial x}\left[v\left(\varrho\epsilon+p+\frac{1}{2}\varrho v^2\right)\right]\ . \tag{7.30}$$

This is a familiar form of the energy equation: On the left we have the change of the density of internal plus kinetic energy; on the right there is the negative divergence of the flow of these two forms of energy, plus the work that must be done due to the presence of a pressure gradient and because the gas is compressible (notice that $h=\epsilon+p/\varrho$ is the specific enthalpy).

Equation (7.30) is treated in an analogous manner as the equation of continuity and the equation of motion. This leads to the third of the *Rankine–Hugoniot relations*,

$$\left[\left[\epsilon+\frac{p}{\varrho}+\frac{1}{2}v^2\right]\right]=0\ , \tag{7.31}$$

where relation (7.27) has been used

The Rankine–Hugoniot relations are three equations for 6 quantities, namely v, p, and ϱ on either side of the shock. The specific internal energy ϵ is assumed to be a known function of p and ϱ. In the case of a perfect gas with constant specific heats we have $\epsilon=p/(\gamma-1)\varrho$.

From the first two Rankine–Hugoniot relations, (7.27) and (7.29), we eliminate the velocities:

$$v_4^2=V_4^2\frac{p_4-p_\mathrm{r}}{V_\mathrm{r}-V_4}\ ,\qquad v_\mathrm{r}^2=V_\mathrm{r}^2\frac{p_4-p_\mathrm{r}}{V_\mathrm{r}-V_4}\ , \tag{7.32}$$

where the subscripts $(4,\mathrm{r})$ denote the values on the left and right of the front, respectively, and where $V_i=1/\varrho_i$. Substitution into (7.31) yields a relation for the thermodynamic variables alone,

$$\epsilon_\mathrm{r}-\epsilon_4=\frac{1}{2}(p_4+p_\mathrm{r})(V_4-V_\mathrm{r})\ , \tag{7.33}$$

which is known as a *Hugoniot curve*.

7.2.4 Analytical reference solution

One may use the Rankine–Hugoniot relations to derive an analytical solution for the Riemann problem, piecewise for the diverse regimes separated by the $x_i(t)$ shown in Fig. 7.3. We do this[1] for the variables v, p, and ϱ, where the velocity v will now always be that in the rest frame; the specific entropy may then be evaluated by the first of (7.20).

[1]The present text takes some advantage of the home page of W. Kley.

The values at the left and right ends of the tube, v_l, p_l, ϱ_l and v_r, p_r, ϱ_r, respectively, are assumed to be known. The values in the left ($x < x_1$) and right ($x > x_4$) intervals are equal to the respective boundary values. In the actual experiment described above, we had $v_\mathrm{l} = v_\mathrm{r} = 0$.

Positions of the separating interfaces

The three equations of the Riemann problem remain invariant under a transformation $x' = fx$, $t' = ft$, with an arbitrary scaling factor f. This means that the solution is a *similarity solution*: it remains similar to itself at all times. Therefore the positions of the separating interfaces between the diverse regimes must be linear functions of time. If we assume that the experiment starts at $t = 0$ with the pressure discontinuity at $x = 0$, all interfaces will move as

$$x_i(t) = v_i^{(\mathrm{int})} t \ , \tag{7.34}$$

such as illustrated in Fig. 7.3. We will determine the interface velocities $v_i^{(\mathrm{int})}$ in the following.

The boundaries of the expansion zone, x_1 and x_2, move backwards with the local speed of sound, but adverse to the local velocity of the gas flow, i.e.,

$$x_1 = (v_\mathrm{l} - c_\mathrm{sl})t \ , \qquad x_2 = (v_3 - c_\mathrm{s3})t \ . \tag{7.35}$$

The contact discontinuity is a consequence of the fact that the initial values p_l, ϱ_l and p_r, ϱ_r are fixed in an arbitrary manner. The dynamics produces pressure equilibrium $p_3 = p_4$, while a density jump remains. This density jump is advected with the flow velocity, so that

$$x_3 = v_3 t \ . \tag{7.36}$$

In order to determine the shock position, x_4, we recall that the velocities that occur in the Rankine–Hugoniot relations are relative to the moving shock. Hence we may solve relations (7.27) and (7.29) for the velocity v_s of the shock (in the rest frame),

$$v_\mathrm{s} = v_\mathrm{r} + \frac{p_4 - p_\mathrm{r}}{\varrho_\mathrm{r}(v_4 - v_\mathrm{r})} \ , \tag{7.37}$$

where the yet unknown variables v_4 and p_4 will be determined presently. The shock position is at

$$x_4 = v_\mathrm{s} t \ . \tag{7.38}$$

The post-shock regime

Next we derive two equations for the post-shock quantities v_4 and p_4. From Eqs. (7.32) we obtain

$$v_4 - v_\mathrm{r} = [(p_4 - p_\mathrm{r})(V_\mathrm{r} - V_4)]^{1/2} \tag{7.39}$$

(notice that v_4 and v_r in Eqs. (7.32) were the velocities relative to the shock front and are negative here; the sign of the square root is consistent with $v_4 > v_\mathrm{r}$). From (7.33) we eliminate $\epsilon_i = p_i/(\gamma{-}1)\varrho_i$, $i = 4, \mathrm{r}$, which yields

$$V_4 = V_\mathrm{r} \frac{(\gamma{-}1)p_4 + (\gamma{+}1)p_\mathrm{r}}{(\gamma{+}1)p_4 + (\gamma{-}1)p_\mathrm{r}} \ . \tag{7.40}$$

We use this to eliminate V_4 from (7.39) and obtain

$$v_4 - v_\mathrm{r} = (p_4 - p_\mathrm{r}) \left\{ \left[(\gamma{+}1)p_4 + (\gamma{-}1)p_\mathrm{r} \right] \frac{\varrho_\mathrm{r}}{2} \right\}^{-1/2} \ . \tag{7.41}$$

This is one equation for v_4 and p_4. A second equation can be derived from the connection to the left-most region, via the rarefaction wave. There, the *Riemann invariant*

$$\Gamma_+ = v + \int c_\mathrm{s} \frac{d\varrho}{\varrho} \tag{7.42}$$

along the *characteristic curves*[2] $x = (v - c_\mathrm{s})t$ yields $v + 2c_\mathrm{s}/(\gamma{-}1) = \text{const.}$ [Use $c_\mathrm{s}^2 = p\gamma/\varrho$ and $p \propto \varrho^\gamma$; hence $d\varrho/\varrho = (dp/p)/\gamma = 2(dc_\mathrm{s}/c_\mathrm{s})/(\gamma{-}1)$]. As all those characteristic curves emerge from the origin, the invariant must have the same value in all regions, and we have

$$\Gamma_+ = v_\mathrm{l} + \frac{2c_\mathrm{sl}}{\gamma - 1} = v_2 + \frac{2c_\mathrm{s2}}{\gamma - 1} = v_3 + \frac{2c_\mathrm{s3}}{\gamma - 1} \tag{7.43}$$

for all $(x,\, t)$ in the rarefaction wave and the two adjacent zones. Especially, with

$$\frac{c_\mathrm{s3}}{c_\mathrm{sl}} = \left(\frac{p_3}{p_\mathrm{l}} \right)^{\frac{\gamma-1}{2\gamma}} \ , \qquad p_3 = p_4 \ , \qquad v_3 = v_4 \ , \tag{7.44}$$

we find the desired second equation for v_4 and p_4:

$$v_4 = v_\mathrm{l} + \frac{2c_\mathrm{sl}}{\gamma - 1} \left[1 - \left(\frac{p_4}{p_\mathrm{l}} \right)^{\frac{\gamma-1}{2\gamma}} \right] \ . \tag{7.45}$$

Once v_4 and p_4 are known,[3] the shock velocity (7.37) can be evaluated, and from the first Rankine–Hugoniot relation, Eq. (7.27), one obtains the post-shock density

$$\varrho_4 = \varrho_\mathrm{r} \frac{v_\mathrm{r} - v_\mathrm{s}}{v_4 - v_\mathrm{s}} \ . \tag{7.46}$$

[2] For an introduction to characteristics and Riemann invariants see, e.g., Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic Press 1966).

[3] To solve the set of equations (7.41), (7.45) in practise, one can simply iterate, using Eq. (7.41) in the form $p_4 = p_r + (v_4{-}v_r)\sqrt{\ldots}$ to update p_4, and then Eq. (7.45) to improve v_4.

The zones to the left of the contact discontinuity

The remaining unknown functions are readily obtained. Between the expansion zone and the contact discontinuity, we have $p_3 = p_4$, $v_3 = v_4$ and

$$\varrho_3 = \varrho_\mathrm{l} \left(\frac{p_3}{p_\mathrm{l}} \right)^{1/\gamma} . \tag{7.47}$$

Within the expansion zone the velocity follows a linear variation:

$$v_2(x,t) = v_\mathrm{l} \frac{x_2(t) - x}{x_2(t) - x_1(t)} + v_3 \frac{x - x_1(t)}{x_2(t) - x_1(t)} . \tag{7.48}$$

The pressure in this region obeys — cf. the derivation of (7.45) above —

$$p_2 = p_\mathrm{l} \left[1 - \frac{\gamma - 1}{2 c_\mathrm{sl}} (v_2 - v_\mathrm{l}) \right]^{\frac{2\gamma}{\gamma - 1}} , \tag{7.49}$$

and the density is

$$\varrho_2 = \varrho_\mathrm{l} \left(\frac{p_2}{p_\mathrm{l}} \right)^{1/\gamma} . \tag{7.50}$$

This completes the construction of the similarity solution to the Riemann problem. The sound velocities $c_{\mathrm{s,i}}$ that appear in the formulas of this section are obtained from the respective pressure and density values.

Bibliography

Hoffman, J. D. (1992). *Numerical Methods for Engineers and Scientists*, McGraw-Hill, New York.

Lele, S. K. (1992). Compact finite difference schemes with spectral-like resolution, *J. Comput. Phys.* **103**: 16–42.

Parker, E. N. (1958). Dynamics of the interplanetary gas and magnetic fields, *Astrophys. J.* **128**: 664–676.

Pozrikidis, C. (1998). *Numerical Computation in Science and Engineering*, Oxford University Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, second edn, Cambridge University Press, Cambridge.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1996). *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, second edn, Cambridge University Press, Cambridge.

Simó, C. (2002). Dynamical properties of the figure eight solution of the three-body problem, *in* A. Chenciner, R. Cushman, C. Robinson and Z. Xia (eds), *Celestial Mechanics.*, Vol. 292 of *Contemporary Mathematics*, American Mathematical Society. [Available at `http://www.maia.ub.es/dsg/2001/index.html`].

Stoer, J. and Bulirsch, R. (1993). *Introduction to numerical analysis*, Vol. 12 of *Texts in applied mathematics*, second edn, Springer, New York.

Index