A First Course in Magnetohydrodynamics

This text introduces readers to magnetohydrodynamics (MHD), the physics of ionised fluids. Traditionally, MHD is taught as part of a graduate curriculum in plasma physics. By contrast, this text – one of a very few – teaches MHD exclusively from a fluid dynamics perspective, making it uniquely accessible to senior undergraduate students. Part I of the text uses the MHD Riemann problem as a focus to introduce the fundamentals of MHD: Alfvéns theorem, waves, shocks, rarefaction fans, and so on. Part II builds upon this with presentations of broader areas of MHD: fluid instabilities, viscid hydrodynamics, steady-state MHD, and non-ideal MHD. Throughout the text, more than 125 problems and several projects (with solutions available to instructors) reinforce the main ideas. In addition, large-font lesson plans for a "flipped-style" class are available free of charge to instructors who use this text as required reading for their course. This book is suitable for advanced undergraduate and beginning graduate students of physics, requiring no previous knowledge of fluid dynamics or plasma physics.

David Clarke is a retired professor of astronomy and physics from Saint Mary's University in Halifax, Nova Scotia. Over his thirty-year career, he has taught numerous courses in physics and astronomy at the undergraduate and graduate levels, including courses in fluid dynamics and MHD that inspired this text. He is co-developer of the original *ZEUS* MHD code and currently the primary developer of *ZEUS-3D* that he uses for his research in astrophysical jets and has made available open source to hundreds of investigators worldwide.

Professor Hannes Olof Gösta Alfvén (1908–1995), father of magnetohydrodynamics, 1970 Nobel Prize laureate for physics. Portrait created in 1972 by Benno Movin-Hermes (1902–1977) using a *three-colour foil method* developed by the artist. Reproduced with permission from I. Movin and the Moderna Museet, Stockholm.

A First Course in

Magnetohydrodynamics

David Alan Clarke
Saint Mary's University, Halifax NS Canada

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009381475

DOI: 10.1017/9781009381468

First published 2025

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Clarke, David Alan, 1958– author.
Title: A First Course in Magnetohydrodynamics / David Alan Clarke, Saint Mary's University, Nova Scotia.
Description: Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2025. | Includes bibliographical references and index.
Identifiers: LCCN 2024014429 | ISBN 9781009381475 (hardback) | ISBN 9781009381468 (ebook)
Subjects: LCSH: Magnetohydrodynamics – Textbooks.
Classification: LCC QC718.5.M36 C53 2025 | DDC 538/.6–dc23/eng/20240629
LC record available at https://lccn.loc.gov/2024014429

ISBN 978-1-009-38147-5 Hardback

For Jodi, my life-long love.

Contents

Preface

WHENEVER a university professor stares down the barrel of a new course preparation, the first question invariably asked is *Is there a text?* For the most part in undergraduate physics, the answer is usually *Yes, and plenty to choose from.* But for a senior undergraduate or beginning graduate course in *magnetohydrodynamics* (MHD), the selection is much narrower.

MHD is a relatively new branch of physics. Developed by Hannes Alfvén during the 1940s, it didn't gain wide acceptance among physicists writ large until the late 1950s culminating in Alfvén's Nobel Prize in 1970.[1] As such, MHD is often touted as a "classical afterthought", the only branch of classical physics introduced *after* Quantum Mechanics with many of its fundamentals – the Riemann problem, magneto-rotational instability, and non-ideal effects – still being worked out in the 1990s and early aughts.

Thus, MHD has not had as long a history as other branches of physics in which textbooks could accumulate, particularly at the undergraduate level. Indeed, MHD has largely been considered a graduate-level subject and, because of this, the vast majority of existing texts specialise in areas such as fusion physics, solar physics, and planetary discs, many written for students already with some familiarity of plasma physics or fluid dynamics.

Another extenuating circumstance is MHD is a divided field whose practitioners – largely plasma physicists and fluid dynamicists – approach the subject in two very different ways. From a plasma physicist's point of view (PoV), an MHD system is the isotropic limit of an ensemble of charged particles – a plasma – governed by velocity moments of the *Vlasov-Boltzmann equation*, a 6-D inhomogeneous partial differential equation (PDE) at the heart of plasma physics. From a fluid dynamicist's PoV, an MHD system is never considered as an ensemble of particles, but rather as a continuous medium governed by simple conservation rules that *any* undergraduate physics student can understand. This leads to a hyperbolic set of equations that can be analysed entirely in terms of *waves*. The two approaches couldn't be more different.

After thirty years of teaching graduate and undergraduate (astro)physics, it is my considered opinion that for MHD to be approachable by undergraduates, it needs to be taught from the fluids PoV. Wave mechanics – so fundamental in classical mechanics, electrodynamics, and quantum mechanics – is already ingrained in the mind of a fourth-year student. On the other hand, velocity moments of a six-

[1] On pages 127–128, there's an amusing anecdote on what – or who – changed the physics community's collective mind on MHD, and a link to Anthony Peratt's short biography of Alfvén.

dimensional PDE are not. Further, most texts taking the plasma PoV are focused on laboratory plasmas and fusion physics, and the wave nature of MHD is often overlooked. To my taste, MHD is *the* prototype for teaching and reinforcing *wave mechanics*, and this is precisely the approach I take in this text.

While there are plenty of textbooks on MHD from the plasma PoV, precious few exist from the fluids PoV. In the survey of texts I did as part of my proposal for this text, I found more than 100 books written over the past six decades from the plasma PoV focused on plasma physics with a portion – sometimes substantial – devoted to the "MHD limit". Indeed, a dozen or so of these include MHD in their titles.[2] By contrast, I found *two* texts on MHD written entirely from the fluids PoV and directed to senior undergraduates: Kendall & Plumpton's *Magnetohydrodynamics with Hydrodynamics* (1964); and Galtier's *Introduction to Modern Magnetohydrodynamics* (2016).

This text offers a third. My approach focuses on the *fundamentals* of the subject and teaches MHD for its own sake rather than dwelling on directed applications and current areas of research; these, I argue, are better suited for graduate texts of which there are plenty. I do provide numerous examples from the literature, but these are selected to emphasise certain ideas (*e.g.*, planetary discs with non-ideal MHD, stellar winds with steady-state MHD, astrophysical jets with Bernoulli's principle, *etc.*) and none should distract the reader from the current discussion. Once endowed with the fundamentals, I contend, students can carry these forward to further their study at the graduate level, should they choose.

In keeping with the undergraduate theme, the first part – *1-D MHD in Ten Weeks* – is designed around a single goal: solving the 1-D MHD Riemann problem. I also assert that to understand MHD, one first has to understand ordinary hydrodynamics (HD) which is, after all, just the zero-field limit of MHD. To these ends, Chap. 1 introduces the student to the fundamentals of HD that includes a novel and simple derivation of the three ideal HD equations. Chapter 2 focuses on 1-D applications of HD including sound waves, shocks, bores, and Bernoulli's principle while Chap. 3 develops a semi-analytic solution to the hydrodynamical Riemann problem. In so doing, students learn how the equations of HD lead to a wave equation, and are shown three ways to extract information about hydrodynamical waves: direct solution of the wave equation; normal mode analysis using linear algebra; and via Riemann invariants and their characteristic paths. In my experience, introducing students to these methods – particularly the latter two – for the relatively simple case of HD is critical for them to understand how they apply to the much more complicated MHD case.

The magnetic induction, $\vec{B}$, doesn't appear until Chap. 4 where the ideal induction equation and the Lorentz force are introduced, along with Alfvén's theorem, magnetic helicity, and flux linking. Chapter 5 examines the MHD equations in 1-D to uncover all three types of waves (slow, Alfvén, fast) and all discontinuities (tan-

[2]The most ambitious and a very recent example of this is Goedbloed, Keppens, & Poedts' *Magnetohydrodynamics of Laboratory and Astrophysical Plasmas* (2019). This is a comprehensive tour de force which could support at least three graduate-level courses.

gential, rotational, shocks) including all three shock subtypes (slow, intermediate, fast). Chapter 6 introduces slow and fast rarefaction fans, and then brings it all together to show students how an exact MHD Riemann solver can be assembled. As this is a semi-analytical solution, students learn or have reinforced semi-analytical techniques including Runge–Kutta methods, multivariate secant root finders, methods for maintaining machine accuracy, and the list goes on.

Part I is designed to be completed in twenty-five hours of instruction (ten weeks at most Canadian universities). The four chapters in Part II, *Additional Topics in (M)HD*, are independent from each other, depend only on material from Part I, and give the instructor options to complete the semester. These include (M)HD instabilities in Chap. 7, viscid HD (Navier–Stokes equation) in Chap. 8, steady-state MHD in Chap. 9, and non-ideal MHD in Chap. 10. In the interest of expediency, sections designated as "optional" can be omitted without loss of continuity.

Parts of the text may come across as "mathematically dense"; this is deliberate. As an undergraduate, I always found it frustrating and distracting when I was unable to fill in the large gaps left between lines of logic in the texts my professors chose, and I was not going to produce a text that did the same. That said, the densest parts of the mathematics can largely be skimmed on first read and certainly don't need to be covered in detail in class, as the main results from which the physics is extracted are always boxed within each development. For the student like I was who needs to know how the derivations are done, the gaps between the mathematical steps are small enough that a careful second read should suffice.

More than 125 problems – many exploiting "teaching moments" – and several computer projects are distributed amongst the ten chapters' problem sections, each generally digestible by a senior undergraduate or first-year graduate student. Problems without an asterisk can and should be done in a page or less (and often a few lines), one asterisk indicates a two-page solution, two asterisks indicate a three to four page solution, while three asterisks indicate a more involved problem, generally requiring more than five pages and/or a substantive computer program to solve. A complete solution set including the computer projects is available to the instructor upon request.

The eight appendices are designed to remind students of particularly critical material prerequisite to this text. Students who do not recognise, recall, or know how to use any of this material are encouraged to review the relevant material from previous courses. Following the appendices is a glossary of symbols used throughout the text, a list of references, and finally an extensive index.

While this text assumes no previous knowledge of (M)HD, students should have had second- and third-year courses in mechanics, electrodynamics, and thermodynamics. On the math side, students should be fluent in vector calculus (at the level of App. A), adept at solving differential equations including PDEs such as the wave equation, and thoroughly familiar with linear algebra and, in particular, *eigen*algebra. In addition, some experience in scientific computing (algorithm and code development) would be beneficial.

Finally, an acknowledgement of the biases of the author is in order. While this text includes numerous astrophysical applications, astronomy is by no means a prerequisite, nor is this text designed just for budding astrophysicists. I would like to think *any* physicist interested in learning about fundamental MHD will find this book useful. As for units, I follow the bulk of the physics community (but not astronomy!) and use `mks` exclusively. Lastly, all program listings in this text are in *FORTRAN*77, the *only* computer language – this old programmer would assert – a computational scientist really needs to know!

Like many of my contemporaries, I learned MHD “at the knee of my advisors”, by reading select chapters in certain texts, by going through journal articles, and talking to experts. As a student and post-doctoral fellow, it always struck me as a bit unfair that all other branches of physics seemed to be taught in more systematic and accessible ways – dedicated courses, self-contained textbooks, problem sets at appropriate levels – but somehow not MHD. Granted, MHD doesn’t enjoy the same “critical mass” of students as other areas of physics, and perhaps this inaccessibility is part of the reason why.

This text – almost two decades in the making – is the textbook I wish I had access to forty years ago when I started out in this game. And now as I enter my retirement, it is my profound hope that within these pages, new students of MHD will find a self-contained introduction to the subject that will help launch them into a fascinating, life-long adventure as I have enjoyed!

Each chapter in Part II benefitted from written projects or theses submitted by Saint Mary’s graduate and undergraduate students taking my (M)HD course and/or working with me as a research student in years past. For these efforts, I thank Joel Tanner, Patrick Rogers, Jonathan Ramsey, Nicholas MacDonald, Michael Power, and Christopher MacMackin.

I thank my editors Nicholas Gibbons, Sarah Armstrong, Stephanie Windows, and Jane Chan at Cambridge University Press for their capable and patient guidance of this first-time author. It definitely made my job a lot less daunting! A big thank-you goes to Patricia Langille at Saint Mary’s Patrick Power Library for doing *all* the heavy lifting in getting permissions for the copyrighted material used in this text; she gets the first signed copy! I would also like to acknowledge the academic freedom afforded to me over the past three decades by Saint Mary’s University that made long-term projects like this possible. Thank you all.

This text was typeset using Donald Knuth’s TEX and Leslie Lamport’s LATEX. Many of the figures were created using Xfig developed by Supoj Sutanthavibul, Ken Yap, Brian V. Smith, and others. Figures from *ZEUS-3D* simulations were created using PSPlot developed by Kevin E. Kohler. Countless members of the scientific community are indebted to these people for placing their software into the public domain.

And then there are the magnificent villages of Ménerbes and Saint-Pierre-

Toirac, France. The most difficult sections of this text for me to write were inspired and completed during my long *séjours* there and, other than my home province of Nova Scotia, I can't think of anywhere else I'd rather spend months at a time than in a small French village. So much about France charms me including, of all things, their speed-limit signs! It's not enough to tell you what the speed limit is, the French also feel the need to tell you you're being *reminded* of what the speed limit is! So in homage to my second country, all footnotes throughout the text serving as a "reminder to the reader" are heralded by the French translation *Rappel*.

And finally to my wife Jodi (MEL) to whom this text is dedicated. Words can't express my love and gratitude for sticking by me these nearly forty years and for helping create such a wonderful and supportive home for our family here in Halifax. You're the best!

Any constructive feedback on what works and doesn't work in this text, as well as any error reports, omissions, redundancies, *etc.* are welcome and can be sent to me directly at AfciMHD@gmail.com. Instructors of courses based on this text may download a solution set to the problems and a fully-developed set of course lecture notes from CUP's website, www.cambridge.org/9781009381475.

David Clarke
Halifax, Nova Scotia
www.ap.smu.ca/~dclarke

1 The Fundamentals of Hydrodynamics

Everything flows and nothing abides; everything gives way and nothing stays fixed.

Heraclitus (c. 535–c. 475 BCE)

1.1 Definition of a fluid

THE PHYSICS of hydrodynamics (HD), namely conservation of mass, conservation of energy, and Newton's second law, are all concepts familiar to first-year undergraduate students, though the mathematics to solve the relevant equations is not. Consider an *ensemble of particles* within some volume V, and let these particles interact with each other via elastic collisions. We can let V remain fixed (in which case we allow the particles to collide elastically with the walls of the container too), or we can let V increase or decrease as the particles move apart or come together; it does not matter. If the mass, total energy, and momentum of the ensemble of particles are M, $E_{\rm T}$, and $\vec{S}$ respectively, then we have:

$$\frac{dM}{dt} = 0, \qquad \text{conservation of mass;} \tag{1.1}$$

$$\frac{dE_{\rm T}}{dt} = \sum \mathcal{P}_{\rm app}, \qquad \text{conservation of total energy;} \tag{1.2}$$

$$\frac{d\vec{S}}{dt} = \sum \vec{F}_{\rm ext}, \qquad \text{Newton's second law.} \tag{1.3}$$

Here, $\sum \mathcal{P}_{\rm app}$ is the rate at which work is done (power) by all forces *applied* to the ensemble of particles, and $\sum \vec{F}_{\rm ext}$ are all forces *external* to and acting on the ensemble of particles. Note that the applied forces – normally just collisions from neighbouring ensembles of particles – are typically a subset of the external forces, which include collisions from neighbouring particles *plus* forces arising from gravity, magnetism, radiation, *etc.* This is because in addition to the thermal and kinetic energies, the *total energy*, $E_{\rm T}$, includes gravitational, magnetic, radiative, and possibly other energies as well.

It is how we model the collisional forces from neighbouring ensembles of particles that defines both what constitutes a fluid and how Eq. (1.1)–(1.3) are further developed. Consider a small cube with volume $\Delta V = (\Delta l)^3$ as shown in Fig. 1.1*a*.

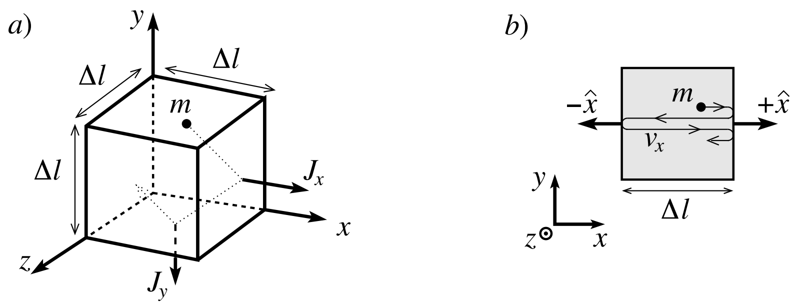


Figure 1.1. *a*) A single particle bounces elastically from the walls of a cube of edge length, Δl, imparting impulses J_x, J_y, *etc.* *b*) An x–y cut through the cube in panel *a* showing one particle whose motion is entirely in the x-direction.

Let the walls of the cube be perfectly reflecting and let there be just one particle inside the cube moving at some speed v in an arbitrary direction.

When the particle collides with the wall, both the particle and cube suffer a change in momentum in a direction normal to the surface of the cube. Moments later, the particle collides with a different wall, and the particle and cube suffer changes in momentum in a direction normal to that wall. A change in momentum is an impulse, J, which when multiplied by the time over which the collision occurs, Δt, constitutes the average force. Thus formally, the "pressure", p, the collision exerts on the wall of the box is this average force divided by the area of the wall:

$$p \sim \frac{J\Delta t}{(\Delta l)^2}.$$

In this scenario, the "pressure" is highly variable in time, and by no means could the "pressure" be construed as isotropic. At a given time, the "pressure" one wall feels will have nothing to do with the "pressures" felt by the other walls.

However, by arbitrarily increasing the number of particles, $\mathcal{N}$, inside our small volume, ΔV, the number of collisions with a given wall, n, occurring in a time Δt will be the same at each wall to within some arbitrarily small variance, Δn. Put another way, averaged over Δt, particle collisions exert the same "pressure" on each wall to within a variance made as small as we please by making $\mathcal{N}$ as large as we please. Thus, we have rendered the particle "pressure" inside the cube *isotropic* because each wall now feels the same force.

There is a contrived exception to this picture. If all the particles were to be placed initially on the mid-plane of the cube and all were launched with the same speed towards one wall of the cube, then it is only with this and the opposite wall that particles would ever collide, and they would do so in a highly ordered, periodic fashion. The remaining four walls would, in principle, never feel any collisions, and thus the "pressure" in the cube would not be isotropic even with $\mathcal{N}$ chosen arbitrarily large. Such a well ordered and well directed ensemble of particles is said to be *streaming* and, as $\mathcal{N}$ is made larger, it becomes increasingly difficult in practice to maintain streaming motion. Small perturbations will eventually cause one particle to collide with another which in turn collide with others, and the ensuing chain reaction quickly reduces the streaming motion to chaos. Isotropic "pressure" (the same "pressure" measured on each of the six walls) is once again the result.

We can now state the key criterion for an ensemble of particles to be treated as a fluid. If there is a sufficient number of particles inside our box (volume element) of dimension Δl so that the motion of particles within the volume element can always be considered isotropic, then the effect of the collisions of particles against the walls of the volume element (which may be rigid walls, or "soft" walls of neighbouring ensembles of particles) is to exert an isotropic "pressure" against all walls. Since isotropy is maintained by particle–particle collisions within the volume element, we may "mathematise" this criterion as,

$$\delta l \ll \Delta l < \mathcal{L}, \tag{1.4}$$

where δl is the mean free path (collision length) of the particles, Δl is the length of one side of our cubic volume element containing an arbitrarily large number of particles, and $\mathcal{L}$ is the smallest length scale of interest in our physical problem. If Ineq. (1.4) holds, we say the ensemble of particles behaves as a *fluid* or a *continuum*. This assumption is an important one; it allows us to treat the applied forces resulting from collisions – which otherwise could be *extremely* difficult to deal with – in a very simple way, namely as an isotropic "pressure".

1.2 A quick review of kinetic theory

To now, I have been enclosing the word *pressure* in quotation marks. This is because I haven't yet made the logical connection between particle collisions (and more specifically, the momentum transferred by particle collisions) and what we commonly think of as *pressure*, such as the *barometric* pressure of the air. So, before we examine how Eq. (1.1)–(1.3) become the equations of hydrodynamics (HD) under the assumption that the ensemble of particles behaves as a fluid (when Ineq. 1.4 is valid), let us review how the "pressure" and the "temperature" of a fluid relate to properties of the ensemble of particles. These ideas form the basis of *kinetic theory*, often exposed to students for the first time in a first-year physics course.[1]

Consider a cube whose edges of length Δl are aligned with the x-, y-, and z-axes of a Cartesian coordinate system, as depicted in Fig. 1.1. Returning to our example in the previous section, suppose a single point particle of mass m moves inside the cube with velocity $v_x\hat{x}$ and collides with the wall whose normal is $+\hat{x}$. If collisions are all elastic, then the particle reflects from the wall with a velocity $-v_x\hat{x}$ and thus suffers a change in momentum of $\Delta S_x = -2mv_x$. Conservation of momentum then demands that an impulse of $+2mv_x$ be imparted against the wall. At a time $\Delta t = 2\Delta l/v_x$ later, the same particle again collides with the wall, imparting another impulse of $+2mv_x$ against it. Thus, the rate at which momentum is delivered to the

[1] For example, Halliday, Resnick, & Walker (2003).

wall by a single particle is given by,

$$\frac{\Delta S_x}{\Delta t} = \frac{2mv_x}{2\Delta l/v_x} = \frac{mv_x^2}{\Delta l} = \langle F\rangle,$$

where $\langle F\rangle$ is the average force felt by the wall. Thus, the average pressure exerted by this one particle, defined as force per unit area, is given by,

$$\langle p\rangle = \frac{\langle F\rangle}{(\Delta l)^2} = \frac{mv_x^2}{V},$$

where $V = (\Delta l)^3$ is the volume of the cube. For $\mathcal{N}$ particles, we simply add over all particles:

$$p \equiv \sum_{i=1}^{\mathcal{N}} \langle p_i\rangle = \sum_{i=1}^{\mathcal{N}} \frac{mv_{x,i}^2}{V} = \frac{m}{V}\sum_{i=1}^{\mathcal{N}} v_{x,i}^2 = \frac{m\mathcal{N}}{V}\langle v_x^2\rangle, \tag{1.5}$$

where each point particle is assumed to have the same mass, m, and where $\langle v_x^2\rangle = \sum v_{x,i}^2/\mathcal{N}$ is the arithmetic mean of the squares of the particle velocities.

For any given particle, $v^2 = v_x^2 + v_y^2 + v_z^2$ and, for large $\mathcal{N}$, one would expect $\langle v_x^2\rangle = \langle v_y^2\rangle = \langle v_z^2\rangle$ since one Cartesian direction shouldn't be favoured over another. Thus,

$$\langle v^2\rangle = \langle v_x^2\rangle + \langle v_y^2\rangle + \langle v_z^2\rangle = 3\langle v_x^2\rangle, \tag{1.6}$$

and Eq. (1.5) becomes,

$$p = \frac{\mathcal{N}mv_{\mathrm{rms}}^2}{3V}, \tag{1.7}$$

where,

$$v_{\mathrm{rms}} \equiv \sqrt{\langle v^2\rangle},$$

is the *root-mean-square* (rms) speed of the particles in the volume V. Comparing Eq. (1.7) with the *ideal gas law*:

$$p = \frac{\mathcal{N}k_{\mathrm{B}}T}{V}, \tag{1.8}$$

(where $k_{\mathrm{B}} = 1.3807 \times 10^{-23}\ \mathrm{J\,K^{-1}}$ is the Boltzmann constant) yields:

$$T = \frac{mv_{\mathrm{rms}}^2}{3k_{\mathrm{B}}} \quad\Rightarrow\quad \frac{3}{2}k_{\mathrm{B}}T = \frac{1}{2}mv_{\mathrm{rms}}^2 = \langle K\rangle, \tag{1.9}$$

where $\langle K\rangle$ is the average kinetic energy per point particle. Thus, while the *pressure*, p, is a measure of the rate at which momentum is transferred from the particles of the fluid (gas) to, for example, the diaphragm of the measuring device (barometer), the *temperature* (or more precisely $3k_{\mathrm{B}}T/2$) is a measure of the average kinetic energy of the particles.

The randomly directed kinetic energy of a system of $\mathcal{N}$ particles is called its *internal energy*, E, and, for the point particles under discussion, is given by,

$$E = \mathcal{N}\langle K\rangle = \frac{3}{2}\mathcal{N}k_{\mathrm{B}}T.$$

The factor $3/2$ is significant and warrants comment. A point particle, as may be found exclusively in a monatomic gas, has three *degrees of freedom* of motion, namely *translation* in each of the three Cartesian directions (Fig. 1.2, left).

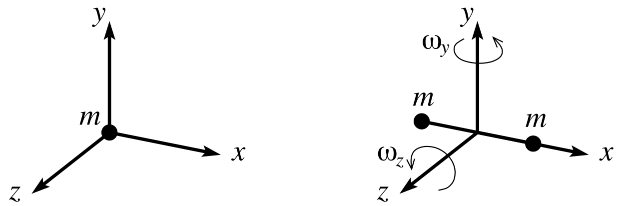


Figure 1.2. A point particle (left) has three degrees of freedom for movement, while a "dumb-bell" (right) has five.

From Eq. (1.6), we have $\langle v_i^2 \rangle = \langle v^2 \rangle/3$ for $i = x, y, z$, and thus to each (translational) degree of freedom we can associate an internal energy $E_i = \mathcal{N} k_\mathrm{B} T/2$, where $E = E_x + E_y + E_z = 3\,E_i$.

Now, a diatomic molecule (essentially two point masses connected by a massless rod) has the same three translational degrees of freedom as a monatomic particle *plus* two *rotational* degrees of freedom, namely rotation about each of the two principle axes orthogonal to its own axis (the x-axis in Fig. 1.2, right), for a total of five degrees of freedom.[2] Note that spinning about the x-axis itself does not constitute a degree of freedom as the moment of inertia about this axis is essentially zero. Because of the *principle of equipartition*,[3] each degree of freedom stores the same amount of kinetic energy, and the internal energy of a diatomic gas must be,

$$E = \frac{5}{2}\mathcal{N} k_\mathrm{B} T.$$

Thus, in general, we write,

$$E = \frac{1}{\gamma - 1}\mathcal{N} k_\mathrm{B} T, \tag{1.10}$$

where $\gamma = 5/3$ for a monatomic gas, $\gamma = 7/5$ for a diatomic gas, and $4/3 \le \gamma < 7/5$ for molecules more complex than diatomic.[4] One can show that $\gamma = C_P/C_V$, the ratio of specific heats of the gas, and that for an adiabatic gas (where heat is neither lost nor gained from the system), $p \propto \rho^\gamma$, where ρ is the *mass density* of the gas.

Dividing Eq. (1.10) by the volume of the sample and using Eq. (1.8) gives an expression for the *internal energy density*, e:

$$e = \frac{E}{V} = \frac{1}{\gamma - 1}\frac{\mathcal{N} k_\mathrm{B} T}{V} = \frac{p}{\gamma - 1}.$$

Thus, an alternate form of the ideal gas law, and the form most frequently used in

---

[2]In principle, there are also two vibrational degrees of freedom which, at "ordinary temperatures", statistical mechanics tells us are insignificant.

[3]Left to their own devices, systems will distribute the available energy equally among all possible ways energy can be stored. Thus, for a large number of diatomic molecules randomly colliding with each other and the walls of their container, one would not expect $m\langle v_x^2 \rangle$ to differ significantly from $m\langle v_y^2 \rangle$ or $m\langle v_z^2 \rangle$ any more than it should differ from $I_y \langle \omega_y^2 \rangle$ or $I_z \langle \omega_z^2 \rangle$, where I_y and I_z are the moments of inertia about the y- and z-axes respectively.

[4]Polyatomic molecules are significantly more complex than diatomic molecules, and the full power of statistical mechanics along with a tensor treatment of its moment of inertia are required to explain the value of γ for any individual molecule.

hydrodynamics, is,

$$\boxed{p = (\gamma - 1)\, e,} \tag{1.11}$$

which states that the rate at which momentum is transferred via collisions is proportional to the average kinetic energy density (*i.e.*, per unit volume) of the random particle motion.

Possibly the second most frequently used form of the ideal gas law in hydrodynamics is,

$$p = \frac{\rho k_{\mathrm{B}} T}{m}, \tag{1.12}$$

which follows directly from Eq. (1.8) noting that $\rho = \mathcal{N}m/V$. Finally, from Eq. (1.9) [and replacing the '3' with $2/(\gamma - 1)$], we find:

$$v_{\mathrm{rms}} = \sqrt{\frac{2k_{\mathrm{B}}T}{(\gamma - 1)\, m}} = \sqrt{\frac{2p}{(\gamma - 1)\, \rho}}. \tag{1.13}$$

Thus, the rms speed goes as the square root of the temperature. We shall encounter another characteristic speed of the gas proportional to the square root of the temperature in §2.1.1, namely the sound speed, c_{s}. Indeed, c_{s} and v_{rms} arise from essentially the same physics, as will be explained when the sound speed is properly introduced.

1.3 The equations of ideal hydrodynamics

In hydrodynamics, the adjective *ideal* means that internal dissipative forces such as viscosity are ignored. A fluid without (with) viscosity is said to be *inviscid* (*viscid*). In this chapter, our discussion is exclusively restricted to inviscid flow. Viscid flow is more the realm of terrestrial HD (though there are important applications for astrophysical fluids as well), and is covered in some depth in Chap. 8.

We begin our discussion by defining the adjectives *extensive* and *intensive*. Variables such as mass, volume, and energy which are *proportional* to the amount of substance being measured are *extensive* quantities, while mass density (often just referred to as density), energy density, and temperature are *independent* of the amount of substance being studied and are examples of *intensive* quantities.

To give a precise relationship between extensive and intensive quantities, consider a small sample of substance with volume ΔV. For every extensive quantity, $Q(V, t)$, of that sample, we can define a corresponding intensive quantity, $q(\vec{r}, t)$, such that,

$$q(\vec{r}, t) = \lim_{\Delta V \to 0} \frac{\Delta Q(V, t)}{\Delta V} = \frac{\partial Q(V, t)}{\partial V}. \tag{1.14}$$

This is a *microscopic* description of the system; q may well change from point to point. A *macroscopic* description of the system can be obtained by integrating Eq. (1.14) over a finite volume, V, to recover Q:

$$Q(V, t) = \int_V q(\vec{r}, t)\, dV. \tag{1.15}$$

Note that Eq. (1.15) requires that q be an *integrable* function of the coordinates over the volume V, and thus q can be discontinuous and have poles of order less than unity. On the other hand, Eq. (1.14) requires that Q be a *differentiable* function of V, and thus it must be both continuous and free from any poles of any order. Evidently, differentiability is a more restrictive requirement than integrability, and this observation will have important consequences as we develop the theory further.

We're now ready to introduce and prove a theorem that provides a particularly simple way to derive the equations of hydrodynamics from the conservation laws of Eq. (1.1)–(1.3).

Theorem 1.1. Theorem of hydrodynamics.[5] *If the time dependence of an extensive quantity, Q, is given by:*

$$\frac{dQ}{dt} = \Sigma, \tag{1.16}$$

where Σ represents the possibly time-dependent "source terms" (reasons for Q not being "conserved"), then the evolution equation for the corresponding intensive quantity, $q(\vec{r},t)$, is given by,

$$\frac{\partial q}{\partial t} + \nabla\cdot(q\vec{v}) = \sigma, \tag{1.17}$$

where $\vec{v} = d\vec{r}/dt$, $Q = \int_V q\,dV$, $\Sigma = \int_V \sigma dV$, and where the product $q\vec{v}$ must be a differentiable *function of the coordinates.*

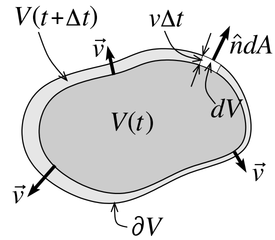
Proof:

$$\frac{dQ}{dt} = \Sigma \quad \Rightarrow \quad \frac{d}{dt}\int_V q\,dV = \int_V \sigma dV,$$

where, in general, the volume element $V = V(t)$ also varies in time. Thus, using the standard definition of the derivative,

$$\begin{aligned}
\frac{d}{dt}\int_{V(t)} q\,dV &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\int_{V(t+\Delta t)} q(\vec{r},t+\Delta t)dV - \int_{V(t)} q(\vec{r},t)dV\right]\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\int_{V(t+\Delta t)-V(t)} q(\vec{r},t+\Delta t)dV\right.\\
&\qquad\qquad \left. + \int_{V(t)} q(\vec{r},t+\Delta t)dV - \int_{V(t)} q(\vec{r},t)dV\right]\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{\Delta V} q(\vec{r},t+\Delta t)dV\\
&\qquad\qquad + \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{V(t)}\big[q(\vec{r},t+\Delta t) - q(\vec{r},t)\big]\,dV,
\end{aligned}$$

[5] This theorem is a variant of *Reynolds' transport theorem*, a volume-integral application of the Leibniz formula for the derivative of an integral.



where, as shown in the inset, performing the volume integral over the difference in volumes, $\Delta V = V(t+\Delta t) - V(t)$, is the same as integrating over the closed surface, ∂V, using a volume differential given by $dV = (\vec{v}\Delta t)\cdot(\hat{n}dA)$. Thus,

$$\begin{aligned}
\frac{d}{dt}\int_{V(t)} q\,dV \;&=\; \lim_{\Delta t\to 0}\frac{1}{\cancel{\Delta t}}\oint_{\partial V} q(\vec{r},t+\Delta t)(\vec{v}\cancel{\Delta t})\cdot(\hat{n}dA) \\
&\qquad + \int_{V(t)} \lim_{\Delta t\to 0}\frac{q(\vec{r},t+\Delta t)-q(\vec{r},t)}{\Delta t}dV \\
&= \oint_{\partial V} q(\vec{r},t)\vec{v}\cdot\hat{n}dA + \int_{V(t)}\frac{\partial q(\vec{r},t)}{\partial t}dV \\
&= \int_{V(t)} \nabla\cdot\big(q(\vec{r},t)\vec{v}\big)dV + \int_{V(t)}\frac{\partial q(\vec{r},t)}{\partial t}dV \quad \text{(Gauss; Eq. A.30)} \\
&= \int_{V(t)}\left(\frac{\partial q(\vec{r},t)}{\partial t} + \nabla\cdot\big(q(\vec{r},t)\vec{v}\big)\right)dV \;=\; \int_{V(t)}\sigma(\vec{r},t)dV \\
&\Rightarrow\quad \int_V\left(\frac{\partial q}{\partial t} + \nabla\cdot(q\vec{v}) - \sigma\right)dV \;=\; 0.
\end{aligned}$$

As this is true for any V, the integrand must be zero, proving the theorem. □

Note that q is not the conserved quantity, Q is (at least to within a known source term, Σ). However, since Q is the volume-integral of q, we'll refer to q as a *volume-conserved quantity*.

The quantity $q\vec{v} \equiv \vec{f}_Q$ is the *advective flux density* of Q whose units are that of Q times $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$; this will require a little unpacking. The *flux*,[6] $\mathcal{F}_Q$, of a vector field, $\vec{f}_Q$, is a measure of how much $\vec{f}_Q$ "passes through" a given surface area with arbitrary normal, $\hat{n}$. Mathematically,

$$\mathcal{F}_Q \;=\; \oint_S \vec{f}_Q\cdot\hat{n}\,dA \qquad \text{or} \qquad \mathcal{F}_Q \;=\; \int_\Sigma \vec{f}_Q\cdot\hat{n}\,dA, \tag{1.18}$$

depending on whether the surface is closed (S) or open (Σ) respectively. Thus, the units of $\vec{f}_Q$ are those of $\mathcal{F}_Q$ per unit area, and $\vec{f}_Q$ can also be interpreted as a *flux density* of $\mathcal{F}_Q$. And so, $\mathcal{F}_Q$ is the flux of $\vec{f}_Q$ while $\vec{f}_Q$ is the flux *density* of $\mathcal{F}_Q$.

An *advective* flux density is more specific to fluid dynamics and refers to some quantity, Q, being *advected* (*i.e.*, transported) by the flow across a surface at a certain rate. Thus, while $\vec{f}_Q$ is the *flux density* of $\mathcal{F}_Q$ with units of $\mathcal{F}_Q$ per unit area, $\vec{f}_Q = q\vec{v}$ is also the *advective* flux density of Q – the volume integral of q – with units of Q per unit area per unit time. It is the "per unit time" part that triggers the adjective *advective*.

Evidently, we have four different types of "fluxes" to keep straight (flux, flux

[6] From the Latin *fluxus* or "flow", this term was introduced to physics by Sir Isaac Newton.

density, advective flux, advective flux density) and the literature seems to blur all four; often you'll find any or all of these terms used interchangeably. In this text, while I maintain the distinction between *flux* and *flux density*, I've chosen to drop the adjective *advective* to simplify the language a bit, relying instead on context. If a particular flux/flux density has a "per unit time" aspect to it, it is an *advective* flux/flux density; otherwise just flux/flux density.

Last point before getting to the equations of HD: Eq. (1.16) is an *integral* equation (Q and Σ both being volume integrals of intensive quantities, q and σ), and thus represents a *global* statement (valid over a finite sample of the fluid) on the conservation of the extensive quantity, Q. On the other hand, Eq. (1.17) is a *differential* equation (often referred to as the *differential form* of Eq. 1.16) and thus represents a *local* statement (valid at a point) on the conservation of Q, involving the corresponding intensive quantity, q. *Global and local forms of an equation are not identical.* Because differential equations require the functions to be differentiable, solutions of the differential form of the equations can be more restrictive than those of the integral form where functions need only be integrable. More on this in §1.5.

Example 1.1. Let $Q = M$, the mass of the sample of fluid. Find the evolution equation for the corresponding intensive quantity, $q = \rho$ (mass density).

Solution: From Eq. (1.1), $\Sigma = 0 \Rightarrow \sigma = 0$, and Theorem 1.1 requires that:

$$\boxed{\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) \;=\; 0.} \tag{1.19}$$

This is the *continuity equation*; the first equation of HD. □

Example 1.2. Let $Q = E_{\rm T}$, the total energy of the fluid sample[7] of mass M:

$$E_{\rm T} \;=\; E + \frac{1}{2}Mv^2 + M\phi,$$

where E is the internal (thermal) energy and ϕ is the gravitational potential. Find the evolution equation for the corresponding intensive quantity, the *total energy density*, namely,

$$e_{\rm T} \;=\; e + \frac{1}{2}\rho v^2 + \rho\phi, \tag{1.20}$$

where once again, e is the internal energy density, whose units $\mathrm{J\,m^{-3}} = \mathrm{N\,m^{-2}}$ are the same as those for pressure, as expected from Eq. (1.11).

Solution: From Eq. (1.2), $\Sigma = \mathcal{P}_{\rm app} \Rightarrow \sigma = p_{\rm app}$, the *applied power density* interpreted as the rate at which work is done on a unit volume of the fluid sample by all *applied* forces. Thus, Theorem 1.1 implies:

$$\frac{\partial e_{\rm T}}{\partial t} + \nabla \cdot (e_{\rm T}\vec{v}) \;=\; p_{\rm app}. \tag{1.21}$$

[7] When we introduce magnetism in Chap. 4, we'll add a magnetic term to $E_{\rm T}$.

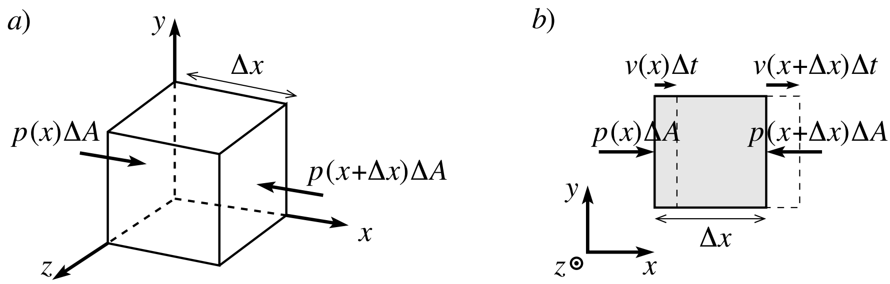


Figure 1.3. *a*) A cube of edge length Δx with external pressure forces acting on the x-faces indicated. *b*) An x–y cut through the cube in panel (*a*) showing both the pressure forces on and motion of the x-faces.

As discussed in §1.1, applied forces are collisions of external particles with the fluid sample. Thus, the applied power is the rate at which work is done on the fluid sample by the external fluid as the former expands or contracts within the latter.

To find an expression for the applied power, $\mathcal{P}_{\rm app}$, consider a small cube of fluid with dimension Δx in the x-direction and cross-sectional area $\Delta A = \Delta V/\Delta x$ (Fig. 1.3). The pressure force exerted on the left face of the cube is $F(x) = +p(x)\Delta A$ and, in time Δt, the left face is displaced by $v_x(x)\Delta t$. Thus, the work done on the left face by the external fluid is $\Delta W_{\rm L} = +p(x)\,v_x(x)\Delta t\,\Delta A$. Similarly, the work done on the right face is $\Delta W_{\rm R} = -p(x+\Delta x)\,v_x(x+\Delta x)\Delta t\,\Delta A$ [since $p(x+\Delta x)$ and $v(x+\Delta x)$ are oppositely directed; Fig. 1.3*b*], and the net work done on the fluid cube is:

$$\begin{aligned}\Delta W \;=\; \Delta W_{\rm L} + \Delta W_{\rm R} \;&=\; p(x)\,v_x(x)\Delta t\,\Delta A - p(x+\Delta x)\,v_x(x+\Delta x)\Delta t\,\Delta A\\ &=\; p(x)\,\delta V(x) - p(x+\Delta x)\,\delta V(x+\Delta x),\end{aligned}$$

where $\delta V(x)$ [$\delta V(x+\Delta x)$] is the small volume change on the left [right] face of the cubic sample of volume ΔV by virtue of the motion of the left [right] face. Because of its form, this work is frequently referred to as the "pdV term".

Dividing ΔW by Δt gives us the applied power,

$$\mathcal{P}_{\rm app} \;=\; \frac{\Delta W}{\Delta t} \;=\; -\Delta A\,\Delta x\,\frac{p(x+\Delta x)\,v_x(x+\Delta x) - p(x)\,v_x(x)}{\Delta x} \;=\; -\Delta V\,\frac{\Delta(pv_x)}{\Delta x},$$

and thus the applied power density is given by:

$$p_{\rm app} \;=\; \frac{\mathcal{P}_{\rm app}}{\Delta V} \;=\; -\frac{\Delta(pv_x)}{\Delta x}.$$

Taking into account similar terms in the y- and z-directions, and letting $\Delta \to \partial$, we have:

$$p_{\rm app} \;=\; -\nabla\cdot(p\,\vec{v}). \qquad (1.22)$$

Substituting Eq. (1.22) into Eq. (1.21) yields:

$$\frac{\partial e_{\rm T}}{\partial t} + \nabla\cdot(e_{\rm T}\vec{v}) \;=\; -\nabla\cdot(p\,\vec{v}),$$

$$\Rightarrow \quad \boxed{\frac{\partial e_{\rm T}}{\partial t} + \nabla\cdot\big((e_{\rm T}+p)\,\vec{v}\big) \;=\; 0,} \tag{1.23}$$

the *total energy equation* and the second equation of ideal HD. □

Example 1.3. Let $Q = \vec{S}$, the *total momentum* of the fluid sample. Find the evolution equation for the corresponding intensive quantity, $q = \vec{s} = \rho\vec{v}$ (the *momentum density*).

Solution: From Eq. (1.3), $\Sigma = \sum \vec{F}_{\rm ext} \Rightarrow \sigma = \sum \vec{f}_{\rm ext}$, the *external force densities.* Thus, Theorem 1.1 requires that:

$$\frac{\partial \vec{s}}{\partial t} + \nabla\cdot(\vec{s}\,\vec{v}) \;=\; \sum \vec{f}_{\rm ext}, \tag{1.24}$$

where the Cartesian representation of the divergence term is:

$$\nabla\cdot(\vec{s}\,\vec{v}) \;=\; \Big(\nabla\cdot(s_x\vec{v}),\nabla\cdot(s_y\vec{v}),\nabla\cdot(s_z\vec{v})\Big).$$

(See §A.4 for other orthogonal coordinate systems.)

For now, we will limit the external force densities to terms arising from pressure gradients and gravity. In Chap. 4, we'll add the Lorentz force, in Chap. 8 viscous stress, and in Chap. 10, drag forces exerted between ions and neutral particles. Starting with the pressure gradient, consider once again the small cube of fluid with edge length Δx and face area ΔA in Fig. 1.3*a*. If the pressure at the left and right sides of the cube are respectively $p(x)$ and $p(x+\Delta x)$, then the net pressure force acting on the cube in the x-direction is given by:

$$F(x+\Delta x) + F(x) \;=\; -p(x+\Delta x)\,\Delta A + p(x)\,\Delta A \;=\; -\frac{\Delta p}{\Delta x}\,\Delta A\,\Delta x \;=\; -\frac{\Delta p}{\Delta x}\,\Delta V.$$

Thus, the pressure force density in the x-direction is:

$$f_x \;=\; \frac{\Delta F_x}{\Delta V} \;=\; -\frac{\Delta p}{\Delta x} \;\to\; -\frac{\partial p}{\partial x} \quad \text{as } \Delta x \to 0.$$

Accounting for all three components,

$$\vec{f}_p \;=\; -\nabla p. \tag{1.25}$$

The gravitational force density, $\vec{f}_\phi$, is even simpler to derive. If the fluid sample has mass ΔM, then the gravitational force on ΔM is $-\Delta M\,\nabla\phi$, where ϕ is the local gravitational potential arising from all external masses, including other regions of fluid and distant or embedded point masses (*e.g.*, stars). Thus, $\vec{f}_\phi$ is given by:

$$\vec{f}_\phi \;=\; -\frac{\Delta M\,\nabla\phi}{\Delta V} \;\to\; -\,\rho\,\nabla\phi \quad \text{as } \Delta V \to 0. \tag{1.26}$$

Substituting both Eq. (1.25) and (1.26) into Eq. (1.24) yields the *momentum equation*, the third and final equation of ideal HD:

$$\boxed{\frac{\partial \vec{s}}{\partial t} + \nabla\cdot(\vec{s}\,\vec{v}) \;=\; -\nabla p - \rho\,\nabla\phi.} \qquad \square \tag{1.27}$$

Summary of §1.3: Equations (1.19), (1.23), and (1.27) constitute two scalar equations and one vector equation which, when combined with Eq. (1.11), (1.20), and $\vec{s} = \rho\vec{v}$ (the *constitutive equations*), provide a closed system of equations for the fluid *flow variables*, namely the volume-conserved quantities ρ, $\vec{s}$, and e_T. This suite of equations comprises our first set of equations of ideal hydrodynamics:

Equation Set 1:

$$
\begin{aligned}
\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\vec{v}) &= 0; && \text{continuity}\\
\frac{\partial e_\mathrm{T}}{\partial t} + \nabla\cdot\big((e_\mathrm{T}+p)\vec{v}\big) &= 0; && \text{total energy equation}\\
\frac{\partial \vec{s}}{\partial t} + \nabla\cdot(\vec{s}\,\vec{v}) &= -\nabla p - \rho\,\nabla\phi; && \text{momentum equation}\\
e_\mathrm{T} &= e + \frac{1}{2}\rho v^2 + \rho\phi; && \text{constitutive equation 1}\\
p &= (\gamma-1)e; && \text{constitutive equation 2}\\
\vec{s} &= \rho\vec{v}. && \text{constitutive equation 3}
\end{aligned}
$$

The gravitational potential, ϕ, is computed by adding up all the potentials of the contributing point masses, and/or by computing the self-gravitational potential of the gas from the density distribution from Poisson's equation:

$$\nabla^2\phi = 4\pi G\rho. \tag{1.28}$$

As a PDE, Poisson's equation is qualitatively different from the equations of hydrodynamics. It has no time derivative, spatial derivatives are second order, and Poisson's equation is an example of an *elliptical* PDE rather than the *hyperbolic* PDEs of HD (App. C). Analytical methods for solving Poisson's equation can be found in any intermediate or advanced text on electrodynamics (*e.g.*, Paris & Hurd, 1969; Lorrain & Corson, 1970; Jackson, 1975 to suggest a few), while numerical treatments can be found in widely available resources such as *Numerical Recipes* (Press *et al.*, 1992). We shall not address such methods in this text.

1.4 The internal energy density

Equation (1.23) governs the evolution of the total energy density, e_T. We can eliminate the need for the first constitutive equation by finding an evolution equation for the internal energy density, e, alone, and our approach shall be via thermodynamics.

The combined first and second law of thermodynamics is:

$$T dS = dE + p\,dV, \tag{1.29}$$

where the only new variable being introduced is S, the total *entropy* of the fluid

6 The MHD Riemann Problem

You don't really understand something until you can compute it.

Michael L. Norman
computational astrophysicist

6.1 Overview

Hands down, the trickiest software I have ever written in my forty years of scientific programming is my program to solve the MHD Riemann problem. It's a venture with zero-divides and near-zero divides around every corner, including the usual and relatively simple-to-solve problems in scalar equations where the denominator gets too close to zero, as well as the much more vexing matrix-vector equations where rows of the Jacobian become zero or near-zero, rendering the matrix equation insoluble or nearly insoluble (*i.e.*, dominated by round-off error). All of these challenges present themselves to those who dare tread forward!

On the plus side, nothing has sealed my own understanding of MHD as has the experience of writing an exact MHD Riemann solver. *Anyone* with serious aspirations of understanding the 1-D MHD problem needs to go through this exercise.

And so let's start with an intuition booster. The precept of the Riemann problem is simple enough. As we did in Chap. 3 and as shown in Fig. 6.1, we start with a left and right state where one state is set completely independently of the other. Before $t = 0$, the two states are separated by an impenetrable diaphragm, **D**, with one state knowing nothing of the other. Then, at $t = 0$ the diaphragm is removed, and suddenly the two states can interact. One state forces its way into the other and yet somehow, at any given time t, one must still be able to get from the left state to the right via a unique set of allowed MHD transitions. The question is, how do we determine these transitions?

Let's approach this by considering a "building block" example, as depicted in Fig. 6.2. I'm thinking of the wooden Brio™[1] train sets my kids and I used to play with when they were little. As shown in panel *a*, suppose there is a vertical gap between A and B that needs to be spanned, and we may do so only with the pieces that come in our set. We'll allow ourselves the latitude of positioning A and B

[1] It should not be lost on the reader that this train set analogy pays homage to one of the seminal papers applying the Riemann problem to computational MHD, namely Brio & Wu (1988).

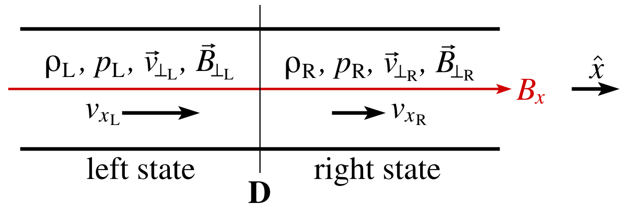


Figure 6.1. Initial set up for the 1-D MHD Riemann problem. At $t = 0$, the diaphragm, **D**, is removed, and the two left- and right-states interact with an arbitrary jump in (possibly) all flow variables at **D** as initial conditions.

horizontally as needed, but their vertical displacement must remain fixed. Suppose further that our Brio$^{\mathrm{TM}}$ pieces – to carry on with the analogy – come in a variety of connectors and colours. Standard Brio$^{\mathrm{TM}}$ pieces are made of maple or birch, and thus light brown with round male–female track connectors at each end. Let's suppose our set comes in four different colours (with just one side of each piece painted), and that all red pieces have a round male and triangular female connector (panel *b*), all green pieces have a triangular male and rectangular female connector (panel *c*), and all blue pieces have a rectangular male and pentagonal female connector (panel *d*). Meanwhile, there is only one black piece, with two male pentagonal connectors (panel *e*). Further, the red and blue pieces come as flats, ramps, and jumps with the ramps and jumps coming in an assortment of heights, the green pieces come in flats and jumps, and the one black piece is flat. As an added wrinkle, all red ramps go down from the male connector while all red jumps go up, opposite to the blue

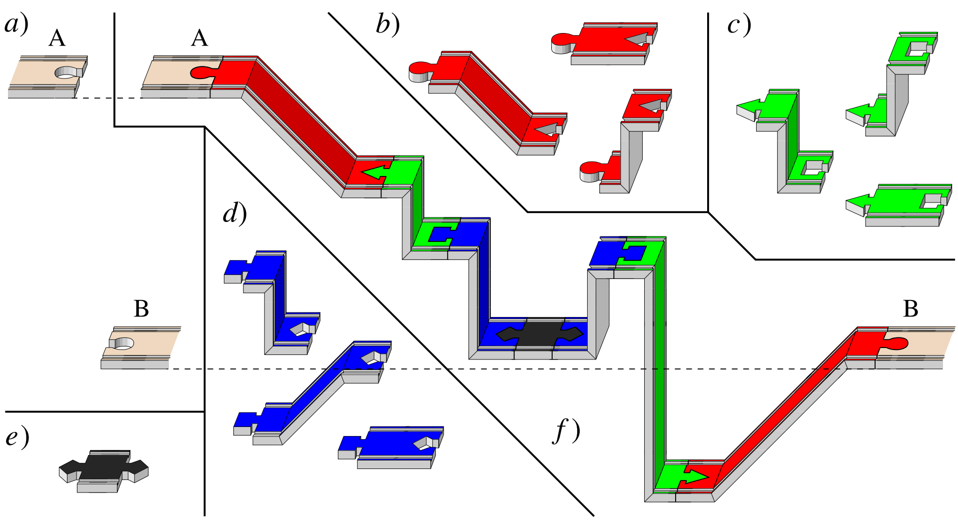


Figure 6.2. The "Brio$^{\mathrm{TM}}$ train track Riemann problem"; see text for description.

pieces where the ramps go up and the jumps go down from the male connector. Green jumps go up or down. Finally, in whatever configuration we come up with to connect A and B, all coloured sides must be face up.

OK, that's what we have and those are the rules. Now let's play!

The first thing to notice is that owing to the shapes of the connectors on each piece, to get from A to B we'll need to arrange the pieces in a specific order. Starting from A, we'll need a red piece, then a green, blue, black, blue, green, and finally a red piece to attach to B. So a first step would be to separate the pieces by colour.

Now comes the harder part where we have to – presumably by trial and error – start fitting pieces to see what combination gets us contiguously from A to B with all connectors joining flatly. It may be that the manufacturer was clever enough to make it so that for any given height difference between A and B, one and only one set of pieces will, in aggregate, span the jump (*e.g.*, panel f in Fig. 6.2). Or perhaps there are numerous solutions, or maybe even none.

The Brio™ game just described is almost a perfect analogy to the 1-D MHD Riemann problem considering just one of the components of $B_\perp$, B_y say. If the vertical distance represents the value of B_y (positive or negative), then the red pieces are the fast waves coming as either rarefaction fans (ramps) or shocks (jumps), the green pieces are the rotational discontinuities, the blue pieces are the slow waves also coming as rarefaction fans or shocks, and the black piece is the contact across which B_y is constant. The real problem, of course, is much more complicated than this since we have not just one gap to fit, but seven – one for each of the variables ρ, p, v_x, v_y, v_z, B_y, and B_z – where each piece chosen for B_y, say, dictates which piece must be used for each of the other variables. Thus, finding one set of pieces to span the B_y gap doesn't necessarily mean the accompanying pieces for ρ, p, *etc.*, will span their gaps. This is starting to look a bit like a 7-D Rubik's cube™!

And so, with that bit of discouragement, let's get started!

6.2 Non-linear MHD waves

From the discussion in §5.3, we're familiar with the MHD discontinuities (contacts, RDs, fast and slow shocks) mentioned in the Brio™ example. All that remains to work out before tackling the MHD Riemann problem directly are the profiles across the fast and slow rarefaction fans (RF). Now, from our discussion in §3.5.3 on the hydrodynamical Riemann problem, we know – at least in principle – how to determine these. Starting from Eq. (3.24), namely,[2]

$$\mathsf{J}_\mathrm{p}|q_\mathrm{p}'\rangle = u_i|q_\mathrm{p}'\rangle,$$

where $|q_\mathrm{p}'\rangle$ is the derivative of the ket of primitive variables with respect to its argument ($\xi_i = x - u_i t$) and J_p is the primitive Jacobian matrix (both defined in Eq.

[2] *Rappel*: This is the *fifth* time we've seen and used this equation, the last time being Eq. (5.55) in §5.2.2.

5.9), we found the seven eigenvalues (characteristic speeds) of J_p to be $u_i = v_x \pm a_\mathrm{f}$, $v_x \pm a_x$, $v_x \pm a_\mathrm{s}$, and v_x, where a_f, a_x, and a_s are the fast, Alfvén, and slow speeds given by Eq. (5.23)–(5.25). Thus, across each wave, $|q'_\mathrm{p}\rangle$ is proportional to one of the right eigenvectors (eigenkets) of J_p, namely $|r_i\rangle$, and, as given in Eq. (3.35),

$$|q'_\mathrm{p}(\xi_i)\rangle = w_i(\xi_i)|r_i\rangle,$$

where w_i is an arbitrary *proportionality* or *scaling* function of the co-moving coordinate, ξ_i. As we did in §3.5.3, define $ds_i = w_i(\xi_i)d\xi_i$ as a differential of a *generalised coordinate*, s_i, that varies from 0 on the upwind side of the i-wave to $s_{i,\mathrm{d}}$ on the downwind side which can be thought of as the "width" or "strength" of the rarefaction fan. Then,

$$\boxed{\frac{d|q_\mathrm{p}(s_i)\rangle}{ds_i} = |r_i\rangle,} \tag{6.1}$$

gives us a set of seven coupled, first-order ODEs which we integrate through to its width, $s_{i,\mathrm{d}}$, to find the profiles (in terms of s_i) of each primitive variable across the i-wave.

And thus we can delay no longer finding the eigenkets of J_p!

6.2.1 Fast and slow eigenkets

The eigenkets of interest here are those associated with eigenvalues $v_x \pm a_\mathrm{f}$ and $v_x \pm a_\mathrm{s}$ respectively, as these describe the fast and slow rarefaction fans. The Alfvén and entropy eigenkets (those associated with eigenvalues $v_x \pm a_x$ and v_x) correspond to the Alfvén and entropy waves, both typically discontinuous in some of the flow variables and thus better handled by the conservative equations (*e.g.*, §5.3.2–5.3.4). Further discussion of these is relegated to Problem 6.1.

Starting with the left-moving fast wave with wave speed $u_1 = v_x - a_\mathrm{f}$, the associated eigenket, $|r_1\rangle = |r_\mathrm{f}^-\rangle$, is found by solving the matrix equation (again, see Eq. 5.9 for J_p),

$$(\mathsf{J}_\mathrm{p} - u_1\mathsf{I})|r_1\rangle = \begin{bmatrix} a_\mathrm{f} & 0 & \rho & 0 & 0 & 0 & 0 \\ 0 & a_\mathrm{f} & \gamma p & 0 & 0 & 0 & 0 \\ 0 & 1/\rho & a_\mathrm{f} & 0 & 0 & B_y/\mu_0\rho & B_z/\mu_0\rho \\ 0 & 0 & 0 & a_\mathrm{f} & 0 & -B_x/\mu_0\rho & 0 \\ 0 & 0 & 0 & 0 & a_\mathrm{f} & 0 & -B_x/\mu_0\rho \\ 0 & 0 & B_y & -B_x & 0 & a_\mathrm{f} & 0 \\ 0 & 0 & B_z & 0 & -B_x & 0 & a_\mathrm{f} \end{bmatrix} \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ r_{14} \\ r_{15} \\ r_{16} \\ r_{17} \end{bmatrix} = 0,$$

which yields seven linear equations, one of which is redundant. So let's try ignoring the third one (if for no other reason, because it has the most number of terms), and write:

$$a_\mathrm{f} r_{11} + \rho r_{13} = 0; \tag{6.2}$$

$$a_\mathrm{f} r_{12} + \gamma p r_{13} = 0; \tag{6.3}$$

$$a_{\rm f} r_{14} - \frac{B_x}{\mu_0 \rho} r_{16} \;=\; 0; \tag{6.4}$$

$$a_{\rm f} r_{15} - \frac{B_x}{\mu_0 \rho} r_{17} \;=\; 0; \tag{6.5}$$

$$B_y r_{13} - B_x r_{14} + a_{\rm f} r_{16} \;=\; 0; \tag{6.6}$$

$$B_z r_{13} - B_x r_{15} + a_{\rm f} r_{17} \;=\; 0. \tag{6.7}$$

As r_{13} appears more often than any other component, let's use that as the pivot (scaling factor). Then, Eq. (6.2) and (6.3) give:

$$r_{11} \;=\; -\frac{\rho}{a_{\rm f}} r_{13}; \qquad r_{12} \;=\; -\frac{\gamma p}{a_{\rm f}} r_{13},$$

and multiplying Eq. (6.4) by $B_x/a_{\rm f}$ and adding Eq. (6.6) gives:

$$-\frac{a_x^2}{a_{\rm f}} r_{16} + B_y r_{13} + a_{\rm f} r_{16} \;=\; 0 \quad\Rightarrow\quad r_{16} \;=\; -\frac{a_{\rm f} B_y}{a_{\rm f}^2 - a_x^2} r_{13}. \tag{6.8}$$

Similarly, Eq. (6.5) and (6.7) yield:

$$r_{17} \;=\; -\frac{a_{\rm f} B_z}{a_{\rm f}^2 - a_x^2} r_{13}. \tag{6.9}$$

Finally, substituting Eq. (6.8) and (6.9) into Eq. (6.4) and (6.5) respectively gives us:

$$r_{14} \;=\; -\frac{B_x}{\mu_0 \rho} \frac{B_y}{a_{\rm f}^2 - a_x^2} r_{13} \qquad \text{and} \qquad r_{15} \;=\; -\frac{B_x}{\mu_0 \rho} \frac{B_z}{a_{\rm f}^2 - a_x^2} r_{13}.$$

Bringing these results together, the "minus fast eigenket" is,

$$|r_{\rm f}^-\rangle \;=\; \psi_{\rm f} \begin{bmatrix} -\rho \\ -\gamma p \\ a_{\rm f} \\ -\dfrac{B_x}{\mu_0 \rho} \dfrac{a_{\rm f}}{a_{\rm f}^2 - a_x^2} B_y \\ -\dfrac{B_x}{\mu_0 \rho} \dfrac{a_{\rm f}}{a_{\rm f}^2 - a_x^2} B_z \\ -\dfrac{a_{\rm f}^2}{a_{\rm f}^2 - a_x^2} B_y \\ -\dfrac{a_{\rm f}^2}{a_{\rm f}^2 - a_x^2} B_z \end{bmatrix}, \tag{6.10}$$

where $\psi_{\rm f} \equiv r_{13}/a_{\rm f}$ is a "scaling factor" which we'll choose for convenience.

Evidently, the "plus fast eigenket" must be identical to Eq. (6.10) with $-a_{\rm f} \to +a_{\rm f}$, and the slow eigenkets are just the fast eigenkets with f $\to$ s. Thus, we have for the four fast and slow eigenkets:

$$|r^{\pm}_{\rm f,s}\rangle \;=\; \psi_{\rm f,s}\begin{bmatrix} -\rho \\ -\gamma p \\ \mp a_{\rm f,s} \\ \pm\dfrac{B_x}{\mu_0\rho}\dfrac{a_{\rm f,s}}{a^2_{\rm f,s}-a^2_x}\vec{B}_\perp \\ -\dfrac{a^2_{\rm f,s}}{a^2_{\rm f,s}-a^2_x}\vec{B}_\perp \end{bmatrix} \;=\; \psi_{\rm f,s}\begin{bmatrix} -\rho \\ -\gamma p \\ \mp a_{\rm f,s} \\ \pm{\rm sgn}(B_x)\,\dfrac{a_x\,a_{\rm f,s}\,a_\perp}{a^2_{\rm f,s}-a^2_x}\,\hat{e}_\perp \\ -\sqrt{\mu_0\rho}\,\dfrac{a^2_{\rm f,s}\,a_\perp}{a^2_{\rm f,s}-a^2_x}\,\hat{e}_\perp \end{bmatrix}, \tag{6.11}$$

where the last four components have been combined into two 2-D vectors $\propto \vec{B}_\perp = (0, B_y, B_z)$, and where $a_x = |B_x|/\sqrt{\mu_0\rho}$, ${\rm sgn}(B_x) = B_x/|B_x|$, $a_\perp = |\vec{B}_\perp|/\sqrt{\mu_0\rho}$, and $\hat{e}_\perp = \vec{B}_\perp/|\vec{B}_\perp|$.

While that may have seemed straight-forward enough, the kicker is the denominator in the components proportional to $\vec{B}_\perp$ ($\hat{e}_\perp$). Because the 1-D MHD equations are *not strictly* hyperbolic, their eigenvalues can, at times, be degenerate. In particular, we've already seen (*e.g.*, Eq. 5.66) that if $\vec{B}_\perp = 0$ and $a_x > c_{\rm s}$, $a_{\rm f} = a_x$ and the components $\propto \hat{e}_\perp$ for the fast eigenkets blow up. Similarly, for $\vec{B}_\perp = 0$ and $a_x < c_{\rm s}$, $a_{\rm s} = a_x$ and the components $\propto \hat{e}_\perp$ for the slow eigenkets blow up.

Oops.

Our salvation are the scaling factors $\psi_{\rm f,s}$, which we choose not so much to normalise $|r^{\pm}_{\rm f,s}\rangle$ (which we can't anyway since the components have different units), but to render all singularities removable. It turns out our choices are rather limited since, in addition, $\psi_{\rm f,s}$ must be chosen such that no eigenket is zeroed out. By what has to be described as a stroke of genius and what we'll spend the next few pages justifying, those introduced by Roe & Balsara (1996) and used by Takahashi & Yamada (2014) are:

$$\psi_{\rm f} \;=\; \sqrt{\frac{c^2_{\rm s}-a^2_{\rm s}}{a^2_{\rm f}-a^2_{\rm s}}} \qquad\text{and}\qquad \psi_{\rm s} \;=\; \sqrt{\frac{a^2_{\rm f}-c^2_{\rm s}}{a^2_{\rm f}-a^2_{\rm s}}}. \tag{6.12}$$

The reader might wonder why the more "obvious" choice of, say, $\psi_{\rm f,s} = a^2_{\rm f,s} - a^2_x$ might not be preferred. However, it doesn't take too long to realise that this would result in $|r^{\pm}_{\rm f,s}\rangle = 0$ for $B_\perp = 0$, which would mean no wave at all.

The mathematics of MHD rarefaction fans is littered with landmines (*i.e.*, zero-divides or, when it comes time to do the programming, *near* zero-divides) which can confound even the most seasoned algebraist. In my experience, the cleanest approach is to express everything in terms of the MHD-alphas, and then to use the various identities among them to eliminate all differences in the denominators where singularities can occur.

So to start, let's recast the four identities among the speeds $a_{\rm f}$, $a_{\rm s}$, a_x, $a_\perp$, and $c_{\rm s}$ as listed in Problem 5.4 (Eq. 5.113–5.116), in terms of the MHD-alphas:

$$a_{\rm f}a_{\rm s} \;=\; c_{\rm s}a_x \quad\Rightarrow\quad \frac{a^2_{\rm f}}{c^2_{\rm s}}\,\frac{a^2_{\rm s}}{c^2_{\rm s}} \;=\; \frac{a^2_x}{c^2_{\rm s}} \quad\Rightarrow\quad \boxed{\alpha_{\rm f}\alpha_{\rm s} \;=\; \alpha_x;} \tag{6.13}$$

$$a^2_{\rm f} + a^2_{\rm s} \;=\; c^2_{\rm s} + a^2_x + a^2_\perp \quad\Rightarrow\quad \boxed{\alpha_{\rm f} + \alpha_{\rm s} \;=\; 1 + \alpha_x + \alpha_\perp;} \tag{6.14}$$

$$\left(a_{\rm f}^2 - c_{\rm s}^2\right)\left(a_{\rm f}^2 - a_x^2\right) \;=\; a_{\rm f}^2 a_\perp^2 \quad\Rightarrow\quad \boxed{(\alpha_{\rm f} - 1)(\alpha_{\rm f} - \alpha_x) \;=\; \alpha_{\rm f}\alpha_\perp;} \tag{6.15}$$

$$\left(c_{\rm s}^2 - a_{\rm s}^2\right)\left(a_x^2 - a_{\rm s}^2\right) \;=\; a_{\rm s}^2 a_\perp^2 \quad\Rightarrow\quad \boxed{(1 - \alpha_{\rm s})(\alpha_x - \alpha_{\rm s}) \;=\; \alpha_{\rm s}\alpha_\perp.} \tag{6.16}$$

Here, $\alpha_{\rm f,s}$ are the fast and slow alphas, first introduced in Problem 5.23 (Eq. 5.124).

Next, examine the fast eigenkets which, from Eq. (6.11), we can write as:

$$|r_{\rm f}^\pm\rangle \;=\; \begin{bmatrix} -\psi_{\rm f}\rho \\ -\psi_{\rm f}\gamma p \\ \mp\psi_{\rm f} a_{\rm f} \\ \pm{\rm sgn}(B_x)\,\chi_{\rm f}\,c_{\rm s}\dfrac{a_x}{a_{\rm f}}\,\hat{e}_\perp \\ -\chi_{\rm f}\,c_{\rm s}\sqrt{\mu_0\rho}\,\hat{e}_\perp \end{bmatrix}, \tag{6.17}$$

where, in terms of the MHD-alphas, $\psi_{\rm f}$ (Eq. 6.12) and $\chi_{\rm f}$ are given by:

$$\psi_{\rm f} \;=\; \sqrt{\frac{1-\alpha_{\rm s}}{\alpha_{\rm f}-\alpha_{\rm s}}}; \qquad \chi_{\rm f} \;=\; \psi_{\rm f}\,\frac{a_{\rm f}^2}{a_{\rm f}^2 - a_x^2}\,\frac{a_\perp}{c_{\rm s}} \;=\; \sqrt{\frac{1-\alpha_{\rm s}}{\alpha_{\rm f}-\alpha_{\rm s}}}\,\frac{\alpha_{\rm f}}{\alpha_{\rm f}-\alpha_x}\,\sqrt{\alpha_\perp}. \tag{6.18}$$

Then, since $c_{\rm s} a_x = a_{\rm f} a_{\rm s}$ (Identity 6.13), and $c_{\rm s}\sqrt{\rho} = \sqrt{\gamma p}$, we may write Eq. (6.17) in its most compact form:

$$|r_{\rm f}^\pm\rangle \;=\; \begin{bmatrix} -\psi_{\rm f}\rho \\ -\psi_{\rm f}\gamma p \\ \mp\psi_{\rm f} a_{\rm f} \\ \pm{\rm sgn}(B_x)\,\chi_{\rm f}\,a_{\rm s}\,\hat{e}_\perp \\ -\chi_{\rm f}\sqrt{\mu_0\gamma p}\,\hat{e}_\perp \end{bmatrix}. \tag{6.19}$$

Now, $\psi_{\rm f}$ and $\chi_{\rm f}$ are much more tightly coupled than Eq. (6.18) appears to suggest. Squaring $\chi_{\rm f}$ and using identity (6.15), we get:

$$\begin{aligned}
\chi_{\rm f}^2 \;&=\; \frac{1-\alpha_{\rm s}}{\alpha_{\rm f}-\alpha_{\rm s}}\,\frac{\alpha_{\rm f}^2}{(\alpha_{\rm f}-\alpha_x)^2}\,\alpha_\perp \;=\; \frac{1-\alpha_{\rm s}}{\alpha_{\rm f}-\alpha_{\rm s}}\,\frac{\cancel{\alpha_{\rm f}^2}(\alpha_{\rm f}-1)^2}{\cancel{\alpha_{\rm f}^2}\alpha_\perp^{\cancel{2}}}\,\cancel{\alpha_\perp} \\
&=\; \frac{\alpha_{\rm f}-1}{\alpha_{\rm f}-\alpha_{\rm s}}\,\frac{(1-\alpha_{\rm s})(\alpha_{\rm f}-1)}{\alpha_\perp} \;=\; \frac{\alpha_{\rm f}-1}{\alpha_{\rm f}-\alpha_{\rm s}}\,\underbrace{(\alpha_{\rm f} - 1 - \underbrace{\alpha_{\rm s}\alpha_{\rm f}}_{\alpha_x} + \alpha_{\rm s})}_{\cancel{\alpha_\perp}}\,\frac{1}{\cancel{\alpha_\perp}},
\end{aligned}$$

using identities (6.13) and (6.14). Thus,

$$\begin{aligned}
\chi_{\rm f}^2 \;&=\; \frac{\alpha_{\rm f}-1}{\alpha_{\rm f}-\alpha_{\rm s}} \;=\; \psi_{\rm s}^2 \quad \text{(Eq. 6.12)} \\
&=\; \frac{\alpha_{\rm f}-\alpha_{\rm s}+\alpha_{\rm s}-1}{\alpha_{\rm f}-\alpha_{\rm s}} \;=\; 1 - \frac{1-\alpha_{\rm s}}{\alpha_{\rm f}-\alpha_{\rm s}} \;=\; 1-\psi_{\rm f}^2.
\end{aligned}$$

Not only is $\chi_{\rm f} = \psi_{\rm s}$, $\chi_{\rm f}$ and $\psi_{\rm f}$ are related to each other in the same way as sine and cosine! Problem 6.2 completes the symmetry of these factors by showing that,

$$\chi_{\rm s} \;\equiv\; \psi_{\rm s}\,\frac{a_{\rm s}^2}{a_x^2 - a_{\rm s}^2}\,\frac{a_\perp}{c_{\rm s}} \;=\; \psi_{\rm f}, \tag{6.20}$$

and thus $\psi_\mathrm{s}^2 + \chi_\mathrm{s}^2 = 1$ as well.

So, let's use these relationships to simplify the notation further by setting $\mu = \psi_\mathrm{f} = \chi_\mathrm{s}$, $\nu = \chi_\mathrm{f} = \psi_\mathrm{s}$ (and thus $\mu^2 + \nu^2 = 1$)[3] and bring the fast and slow eigenkets to our final form:

$$\boxed{|r_\mathrm{f}^\pm\rangle \;=\; \begin{bmatrix} -\mu\rho \\ -\mu\gamma p \\ \mp\mu a_\mathrm{f} \\ \pm\mathrm{sgn}(B_x)\,\nu\, a_\mathrm{s}\,\hat{e}_\perp \\ -\nu\sqrt{\mu_0\gamma p}\,\hat{e}_\perp \end{bmatrix}; \qquad |r_\mathrm{s}^\pm\rangle \;=\; \begin{bmatrix} -\nu\rho \\ -\nu\gamma p \\ \mp\nu a_\mathrm{s} \\ \mp\mathrm{sgn}(B_x)\,\mu\, a_\mathrm{f}\,\hat{e}_\perp \\ \mu\sqrt{\mu_0\gamma p}\,\hat{e}_\perp \end{bmatrix}.} \tag{6.21}$$

It remains, then, to settle on a final form for the scaling factors μ and ν and to demonstrate that their apparent singularities are, in fact, removable. As given by Eq. (6.18), μ (ψ_f) has a singularity (but, as we'll see, removable) at $\alpha_\mathrm{f} = \alpha_\mathrm{s}$ which happens only at the so-called *triple umbilic* where $\alpha_\perp = 0$ and $\alpha_x = 1$ (and thus $\alpha_\mathrm{f} = \alpha_x = \alpha_\mathrm{s} = 1$; that is, the fast, Alfvén, and slow speeds are *triply* degenerate). If, for convenience, we let,

$$\delta_x \;=\; \alpha_x - 1, \tag{6.22}$$

(and thus $\delta_x \gtrless 0$ for $\alpha_x \gtrless 1$), then the singularity in μ occurs when both $\alpha_\perp$ and δ_x are zero.

By necessity, the MHD Riemann solver we'll design will be semi-analytic. Thus, its reliance on a computer means that even a removable singularity will trigger a zero-divide at the triple umbilic, and we must therefore remove it manually. To do this, it is most convenient to express μ in terms of $\alpha_\perp$ and δ_x, those quantities which when simultaneously zero cause the singular behaviour.

Starting with the definition of the slow speed (Eq. 5.23),

$$a_\mathrm{s}^2 \;=\; \frac{1}{2}\big(c_\mathrm{s}^2 + a^2 - D\big) \;=\; \frac{1}{2}\big(c_\mathrm{s}^2 + a_x^2 + a_\perp^2 - D\big), \tag{6.23}$$

where the discriminant, D, is given by,

$$D \;=\; \sqrt{\big(c_\mathrm{s}^2 + a_x^2 + a_\perp^2\big)^2 - 4c_\mathrm{s}^2 a_x^2},$$

we have,

$$\alpha_\mathrm{s} \;=\; \frac{a_\mathrm{s}^2}{c_\mathrm{s}^2} \;=\; \frac{1}{2}\big(1 + \alpha_x + \alpha_\perp - d\big), \tag{6.24}$$

where,

$$d \equiv \frac{D}{c_\mathrm{s}^2} \;=\; \sqrt{\big(1 + \alpha_x + \alpha_\perp\big)^2 - 4\alpha_x} \;=\; \sqrt{\big(\alpha_\perp + \delta_x\big)^2 + 4\alpha_\perp}, \tag{6.25}$$

after a little algebra. This, incidentally, is the most robust form for d one can use for a computer application, since there are no subtractions under the radical. Computers, being of finite precision, can often subtract what ought to be equal values and end up with a small residual of "round-off noise" which can be negative as often as

[3] Don't confuse μ with μ_0!!

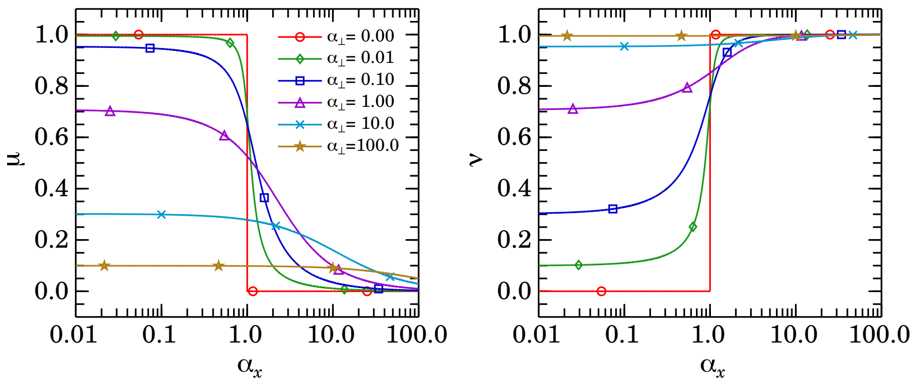


Figure 6.3. Profiles of the fast and slow eigenket "scaling factors", μ (left; Eq. 6.26) and ν (right; Eq. 6.27) as functions of $\alpha_x = \delta_x + 1$, shown for various values of $\alpha_\perp$, including the limiting case of $\alpha_\perp = 0$ (red○) where μ and ν are step functions with the discontinuity at $\alpha_x = 1$ ($\delta_x = 0$), the triple umbilic.

not. Under a radical sign, this would trigger a "floating point exception" or worse, the dreaded "`NaN`" ("not a number") error messages that cause the program to crash, often without any indication of where the first `NaN` occurred! Further, even if the difference is not purely round-off noise, subtracting two nearly equal numbers can yield results significantly less precise than the stated precision of the machine. Conversely, adding positive quantities suffers no such loss of machine accuracy.

Continuing from Eq. (6.24), we have,

$$\alpha_\mathrm{s} = \frac{1}{2}\big(2 + \delta_x + \alpha_\perp - d\big) = 1 + \frac{1}{2}\big(\delta_x + \alpha_\perp - d\big)$$

$$\Rightarrow \quad 1 - \alpha_\mathrm{s} = \frac{1}{2}\big(d - \delta_x - \alpha_\perp\big).$$

Noting that $\alpha_\mathrm{f} - \alpha_\mathrm{s} = d$, we arrive at our final forms for μ and ν:

$$\boxed{\mu^2 = 1 - \nu^2 = \frac{1 - \alpha_\mathrm{s}}{\alpha_\mathrm{f} - \alpha_\mathrm{s}} = \frac{1}{2}\left(1 - \frac{\delta_x + \alpha_\perp}{d}\right);} \tag{6.26}$$

$$\boxed{\nu^2 = 1 - \mu^2 = \frac{\alpha_\mathrm{f} - 1}{\alpha_\mathrm{f} - \alpha_\mathrm{s}} = \frac{1}{2}\left(1 + \frac{\delta_x + \alpha_\perp}{d}\right),} \tag{6.27}$$

with d given in terms of δ_x and $\alpha_\perp$ by Eq. (6.25).

Figure 6.3 shows μ and ν as functions of $\alpha_x = \delta_x + 1$ for various values of $\alpha_\perp$, including the limiting case (red○) where $\alpha_\perp = 0$ ($B_\perp = 0$). The fact that μ^2 (and thus μ) should be a step function when $\alpha_\perp = 0$ is easy to see from Eq. (6.25) and (6.26). Setting $\alpha_\perp = 0$,

$$d = \sqrt{(\delta_x)^2} = |\delta_x|$$

$$\Rightarrow \quad \mu^2 = \frac{1}{2}\left(1 - \frac{\delta_x}{|\delta_x|}\right) = \begin{cases} 1, & \delta_x < 0, \ (\alpha_x < 1); \\ 0, & \delta_x > 0, \ (\alpha_x > 1), \end{cases} \tag{6.28}$$

which is the red○ profile in the left panel of Fig. 6.3. Evidently, the step function is reversed for $\nu^2 = 1 - \mu^2$ (0 for $\alpha_x < 1$, 1 for $\alpha_x > 1$), as the red○ profile in the right panel shows.

The fact that μ and ν remain finite everywhere means that the apparent singularity at the triple umbilic (when $\alpha_{\rm f} = \alpha_{\rm s}$ in Eq. 6.26) is removable; that is, μ^2 remains finite as $\alpha_{\rm f} \to \alpha_{\rm s}$ (when both $\alpha_\perp$ and δ_x are zero). To find that limiting value, we first set $\delta_x = 0$ in Eq. (6.25) and (6.26) to get:

$$\lim_{\delta_x \to 0} \mu^2 = \frac{1}{2}\left(1 - \frac{\alpha_\perp}{\sqrt{\alpha_\perp^2 + 4\alpha_\perp}}\right) = \frac{1}{2}\left(1 - \sqrt{\frac{\alpha_\perp}{\alpha_\perp + 4}}\right) \to \frac{1}{2}, \tag{6.29}$$

as $\alpha_\perp \to 0$. Similarly, $\nu^2 \to \frac{1}{2}$ as δ_x, $\alpha_\perp \to 0$.

This completes the justification of the scaling factors, $\psi_{\rm f,s}$, chosen in Eq. (6.12).

6.2.2 Fast and slow rarefaction fans

Part of the job of the Riemann solver is to integrate Eq. (6.1) using the kets in Eq. (6.21) to find the primitive variable profiles across any rarefaction fan that may be part of the solution. However, without doing the actual integrations, we can determine qualitative properties of the fast and slow fans just by examining the kets as differential changes in the variables.

Assuming for now that $B_x \neq 0$ ($\alpha_x > 0$), let's start by examining the first three components of each ket. Other than the factors μ and ν, these are identical to the purely hydrodynamical kets in Eq. (3.34), with $c_{\rm s}$ replaced with the appropriate wave speed. So, for the moment, let's set $B_\perp$ (and thus $\alpha_\perp$) to zero so that for $\alpha_x < 1$, $\mu = 1$ and $\nu = 0$ (Fig. 6.3). This makes the fast kets *identical* to the hydrodynamical kets for which, in §3.5.3, we concluded that density and pressure *decrease* from the upwind to downwind side of the fan, while the flow speed relative to the upwind state *increases*. A similar comparison between the slow and hydrodynamical fans may be made for $\alpha_x > 1$ where $\nu = 1$ and $\mu = 0$.

Noting from Fig. 6.3 that the primary effect of increasing $\alpha_\perp$ from zero is to round off the discontinuities in μ and ν without changing their monotonic dependence on α_x, we conclude that ρ and p should decrease across *any* RF while the flow speed relative to the upwind state increases. Thus, at least qualitatively, the hydrodynamical variables ρ, p, and v_x behave the same way across a fast and slow fan as they do across a hydrodynamical fan.

A consequence of p dropping across an MHD fan is that $\alpha_x = a_x^2/c_{\rm s}^2 = B_x^2/(\mu_0\gamma p)$ rises. This is an important observation that holds for *any* MHD RF. Note that α_x rises only because p falls; in a 1-D problem such as this, B_x is strictly constant. Thus, one can gain a qualitative feel for how μ and ν vary across a fan – quantities critical to determining the variable profiles from Eq. (6.1) – by scanning across Fig. 6.3 from left to right.

The last component in each ket in Eq. (6.21) governs the profiles of $B_\perp$, and it is here where the properties of the fast and slow fans diverge. For the fast fan, the last component, $-\nu\sqrt{\mu_0\gamma p}$, is negative – just like the first two components governing the

Index