

Max Camenzind

Compact Objects in Astrophysics

White Dwarfs,
Neutron Stars and Black Holes



Springer



ASTRONOMY AND ASTROPHYSICS LIBRARY

Series Editors:

- G. Börner, Garching, Germany
A. Burkert, München, Germany
W. B. Burton, Charlottesville, VA, USA and
Leiden, The Netherlands
M. A. Dopita, Canberra, Australia
A. Eckart, Köln, Germany
T. Encrenaz, Meudon, France
E. K. Grebel, Basel, Switzerland
B. Leibundgut, Garching, Germany
J. Lequeux, Paris, France
A. Maeder, Sauverny, Switzerland
V. Trimble, College Park, MD, and Irvine, CA, USA

Max Camenzind

Compact Objects in Astrophysics

White Dwarfs,
Neutron Stars and Black Holes

With 187 Figures and 25 Tables

Professor Max Camenzind
Landessternwarte Königstuhl
Center for Astronomy Heidelberg (ZAH)
Königstuhl 12
69117 Heidelberg, Germany

Cover picture: Shadow of a black hole: Image of a torus around a rotating black hole seen under an inclination of 60 degrees. Due to Doppler beaming, emission from the approaching side appears strongly enhanced.

Library of Congress Control Number: 2006938305

ISSN 0941-7834

ISBN 978-3-540-25770-7 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting and production: LE-T_EX Jelonek, Schmidt & Vöckler GbR, Leipzig, Germany
Cover design: eStudio Calamar Steinen

Printed on acid-free paper SPIN: 11394105 55/3100/YL - 5 4 3 2 1 0

To Martina, Sven and Tessa,
my friends and coworkers

Preface

In astronomy, a compact star (sometimes called a compact object) is a star that is a white dwarf, a neutron star or a black hole. Our Galaxy is populated by billions of white dwarfs, a few hundred million neutron stars and probably by a few hundred thousand black holes. Of all these objects, only a very tiny fraction has been detected so far by astronomical instruments, just a few thousand white dwarfs, about 2000 neutron stars, and only a few dozen black holes. Of all these objects, only black holes can appreciably grow in mass. It is one of the great successes of the last 15 years that it could be shown that practically every center of galaxies harbors a supermassive black hole with a mass in the range of one million to a few billion solar masses. The visible Universe therefore contains at least 100 billion supermassive black holes. Only about 100,000 of these objects have now been detected as quasars and only about 50 as mass centers of nearby galaxies. Black holes of varying mass are also thought to be the driver behind gamma bursters.

Compact stars form the endpoint of stellar evolution. A star shines and thus loses its nuclear energy reservoir in a finite time. When a star has exhausted all its energy (which is called a stellar death), the gas pressure of the hot interior can no longer support the weight of the star and the star collapses to a denser state – a compact star. One could see the compact stars, such as the white dwarf and the neutron star, as a solid state as opposed to the gaseous interior of all other stars. In contrast to this, the interior of a black hole is very enigmatic. Its surface is formed by a kind of semipermeable membrane forbidding any classical emission from its surface. The very source of the gravitational field of black holes is a kind of curvature singularity, which is hidden behind this membrane. It is expected that quantum effects will smooth these singular mass currents in the center of a rotating Black hole.

A normal star is a fully Newtonian object, in the sense that its gravitational field is a mere solution of the Poisson equation. Gravity of compact objects, on the other hand, must rely on the concepts of space and time. The classical textbook by Shapiro and Teukolsky [15] on the theory of white dwarfs, neutron stars and black holes handles many aspects on these objects. In the last 20 years, however, a great deal of observational data and theoretical insights into the physics of compact objects force us to a more complicated approach for modelling. Just to mention one example: though the Tolman–Oppenheimer–Volkoff equation is still the basis for the calculation of the interior structure of neutron stars, the inclusion of rotation for these objects leads to a nontrivial set of partial differential equations for handling the gravitational field of rapidly rotating neutron stars. For this reason, the author of

this book has decided to base the description of gravity on the general framework which is nowadays used in numerical computations when Einstein's equations are involved.

The concept of this book therefore relies heavily on the concepts of modern gravity. For this reason, Chap. 2 gives an overview of the modern description of gravity. This does not, however, preclude any study of classical textbooks on Einstein's theory of gravity.

Compact objects such as white dwarfs and neutron stars have extremely high densities that cannot be created in terrestrial laboratories and involve phases of matter that are not yet well understood. In these lectures we will work out the associated highly relativistic phenomena theoretically and observationally. One theoretical focus is understanding the interplay between magnetic and thermal processes for strongly magnetic neutron stars. In addition, just like their stellar precursors, many compact objects occur in binary systems. We will study the origin and evolution of compact X-ray binaries using data from RXTE as well as ASCA and ROSAT and other X-ray data. With the successful launch of Chandra and XMM–Newton, X-ray astronomy is in a key position to conduct new high-resolution imaging and spectral studies of compact objects in both binaries and AGN.

High-energy gamma-ray bursts are being detected with regularity now, but their nature remains a mystery. Researchers are actively involved in modelling these bursts and identifying tests and consequences of suggested mechanisms for a wide array of data sets. Cosmic gamma-ray bursts are important for their own intrinsic physics as well as for providing a probe of cosmology. We still do not know the nature of the tremendous explosions that in about one minute release a few percent of a solar mass of rest energy in the form of gamma-rays. However, several clues point to an association with the explosions of massive stars, and current models assume that a gamma-ray burst is triggered by the formation of a black hole.

The study of compact objects probes physics at extreme conditions of density, temperature, and magnetic fields. The mass–radius relation for neutron stars, for example, probes the equation of state at supranuclear densities and may reveal in the future the existence of quark matter in one of the color-superconducting phases. Accurate neutron star masses can be measured for some binaries, especially those including radio pulsars; measuring radii is more difficult, but may be possible through studies of gravitational redshifts, neutron star cooling or the dynamics of gas near the innermost stable circular orbit predicted by general relativity.

Different models for the composition and equation of state of neutron-star matter produce neutron-star models with different properties which might then be detected in observations. Particularly important in this context is the possibility of constraining the form of particle interactions in high-density matter or of finding evidence for the occurrence of phase transitions in the stellar interior or of exotic states of matter (strange stars being an extreme example).

Important stellar evolution questions are being addressed concerning the evolutionary pathways to each of the endpoints for compact objects. Binary star systems can undergo complex mass transfer evolutionary phases. In particular, considerable insight has been gained into how close binary systems containing compact ob-

jects are formed from primordial binaries in the Galaxy and via dynamical capture processes in globular star clusters. Once an accreting compact binary forms, many questions remain about the accretion process itself. For example, largely through observational work conducted with the Rossi X-ray Timing Explorer Satellite (RXTE), astronomers have found that accreting neutron stars often flicker quasiperiodically at frequencies ranging from a few hertz to more than one kilohertz. The cause of this flickering is poorly understood, but may involve effects of strong field gravity in the accretion disk or oscillations of the neutron star.

One exciting fact is that compact objects offer the ultimate strong-field tests of general relativity through the gravitational radiation emitted when black holes form. The recent detection of a double pulsar system opens up a new window on testing relativistic gravity by using compact objects. Together with black holes, these neutron stars will provide the deepest insight into the structure of relativistic gravity. These systems are sources of gravitational waves. The existence and ubiquity of gravitational waves is an unambiguous prediction of Einstein's theory of general relativity. Although gravitational radiation has not yet been unambiguously and directly detected, there is already significant indirect evidence for its existence. Most notably, observations of binary pulsars, which are thought to consist of two neutron stars orbiting rather tightly and rapidly around each other, have revealed a gradual in-spiral at exactly the rate which would be predicted by general relativity.

Heidelberg, December 2006

Max Camenzind

Acknowledgements

Many of my colleagues and former students helped me in improving the original manuscript, which originated from lectures given at the University of Heidelberg. Many thanks go to Stefan Appl, Ramon Khanna, Christian Fendt, Jochen Peitz, Stefan Spindeldreher, Martin Krause, José Gracias, Andreas Müller, Ahmad Hujeirat, Matthias Stute, Matthias Vigelius, Volker Gaibler and Steffen Brinkmann. I am in particular indebted to Tom Mädler who went through all the jungle of the 3+1 split of Einstein's equations. Andreas Bauswein worked out many details about rotating neutron star structure. I am also deeply indebted to my wife Martina who was always patient with me when my brain was full of formulae.

Contents

1	Compact Objects in Astrophysics	1
1.1	Why is Newtonian Gravity Obsolete?	1
1.2	Einstein was Skeptical about the Existence of Black Holes	3
1.3	Subrahmanyan Chandrasekhar and Compact Objects	4
1.4	Classes of Compact Objects	5
1.4.1	White Dwarfs and Neutron Stars	8
1.4.2	Compact X-Ray Sources	9
1.4.3	Radio Pulsars	11
1.5	Supermassive Black Holes in Galactic Centers	16
1.6	Gamma-Ray Bursters	19
	Problems	25
2	Gravity of Compact Objects	27
2.1	Geometric Concepts and General Relativity	27
2.2	The Basic Principles of General Relativity	29
2.2.1	Einstein's Equivalence Principle and Metricity	29
2.2.2	Metric Theories of Gravity	33
2.3	Basic Calculus on Manifolds	37
2.3.1	Tensors and Forms on Manifolds	37
2.3.2	The Metric Field and Pseudo-Riemannian Manifolds	42
2.3.3	The Calculus of Forms on Lorentzian Manifolds	44
2.4	Affine Connection and Covariant Derivative	47
2.4.1	Affine Connection	47
2.4.2	Covariant Derivative of Vector Fields	47
2.4.3	Covariant Derivative for Tensor Fields	48
2.4.4	Parallel Transport and Metric Connection	50
2.4.5	Metric Connection	52
2.4.6	Divergence of Vector Fields	55
2.5	Curvature of Pseudo-Riemannian Manifolds	56
2.5.1	Mathematical Definition of Torsion and Curvature	57
2.5.2	Bianchi Identities for Metric Connection	58
2.5.3	Ricci, Weyl and Einstein Tensor	60
2.5.4	Cartan's Structure Equations	61
2.6	Gravity is a Lorentzian Connection on Spacetime	65
2.6.1	The Four Key Principles of General Relativity	65

2.6.2	The Hilbert Action and Einstein's Field Equations	68
2.6.3	On the Cosmological Constant	69
2.6.4	Limits of General Relativity	71
2.7	Gravitational Waves	73
2.7.1	The Geodesic Deviation – Relativistic Tidal Forces	73
2.7.2	Gravity Wave Experiments	74
2.7.3	The Nature of Gravitational Waves	76
2.7.4	Degrees of Freedom	79
2.7.5	Gravitational Wave Solutions	83
2.7.6	The Quadrupole Formula	87
2.8	3+1 Split of Einstein's Equations	91
2.8.1	Induced Spatial Metric and Extrinsic Curvature	92
2.8.2	Hypersurface Embedding	93
2.8.3	Split of Affine Connection and Curvature	95
2.8.4	Split of Einstein's Equations	98
2.8.5	Black Hole Simulations and Gravitational Waves	100
	Problems	101
3	Matter Models for Compact Objects	105
3.1	General Relativistic Hydrodynamics	105
3.1.1	Relativistic Plasma Equations	106
3.1.2	On Numerics of Hydrodynamics	110
3.2	The Boltzmann Equation in GR	113
3.2.1	The Geodesics Spray on the Cotangent Bundle	113
3.2.2	Particle Number Current and Energy–Momentum Tensor	116
3.2.3	The Relativistic Boltzmann Equation	117
3.2.4	Liouville Operator in 3+1 Split	118
3.2.5	Transformation into the Local Rest Frame	119
	Problems	120
4	Relativistic Stellar Structure	123
4.1	Spacetime of Relativistic Stars	123
4.2	Derivation of the TOV Equations	125
4.2.1	The Curvature of Static Spacetimes	125
4.2.2	Matter in the Interior	127
4.2.3	The Exterior Schwarzschild Solution	130
4.2.4	Stable Branches for Degenerate Stars	131
4.2.5	Metric for Relativistic Stars	131
4.3	A Variational Principle for the Stellar Structure	132
	Problems	134

5	White Dwarfs	137
5.1	Observations of Isolated White Dwarfs	138
5.1.1	Sirius B	138
5.1.2	Field White Dwarfs and Classification	139
5.1.3	White Dwarfs in Globular Clusters	143
5.1.4	Magnetic White Dwarfs	143
5.1.5	Ultracool White Dwarfs as Cosmochronometers	145
5.2	What is Inside a White Dwarf?	151
5.3	Equation of State below the Neutron Drip Density	153
5.4	Structure of White Dwarfs and the Chandrasekhar Mass	159
5.4.1	Polytropic Approximation	160
5.4.2	Beyond the Chandrasekhar Treatment	162
5.4.3	Comparison with Observations	162
5.5	The Relativistic Instability of White Dwarf Stars	167
5.5.1	Necessary Condition for Stability	168
5.5.2	The Total Energy in the Post-Newtonian Limit	169
5.5.3	GR White Dwarf Instability	171
5.6	Cooling White Dwarfs	174
5.6.1	Structure of the Surface Layers	175
5.6.2	Cooling Curves and Crystallization	177
5.6.3	Testing WD Crystallization Theory	179
5.7	White Dwarfs in Binary Systems	180
	Problems	185
6	Neutron Stars	187
6.1	The Structure of a Neutron Star	188
6.2	Equations of State beyond Neutron Drip	189
6.2.1	From Neutron Drip to Saturation	190
6.2.2	Nuclear EoS for Dense Neutron Matter	199
6.2.3	Relativistic Mean Field Theory above Saturation	206
6.2.4	Analytical Fits to EoS	216
6.3	Neutron Star Models	219
6.3.1	Hadronic Models	219
6.3.2	Quark Matter Cores	224
6.3.3	Grand Canonical Potential for Quark Matter	231
6.3.4	Strange Quark Stars	241
6.3.5	The Structure of Massive Neutron Stars	242
6.4	Neutron Stars in Close Binary Systems	244
6.4.1	Post-Newtonian Potentials for Many-Body Systems	244
6.4.2	Periastron Shift in Two-Body Systems	248
6.4.3	The Shapiro Time Delay in a Binary System	250
6.4.4	Decay of Binary Orbits due to Gravitational Radiation	251
6.5	Masses of Neutron Stars from Radio Pulsar Timing	255
6.5.1	What is Pulsar Timing?	255
6.5.2	The Timing Formula	259

6.5.3	Timing of the Binary System PSR B1913+16	263
6.5.4	Masses of Companion Stars	264
6.5.5	The Double Pulsar System PSR 0737-3039A+B	265
6.6	Neutron Stars in our Galaxy	269
6.6.1	100 Million Neutron Stars in the Galaxy	269
6.6.2	Thermal Emission from Isolated Neutron Stars	272
6.6.3	Rotation-Powered Pulsars	284
6.6.4	Accretion-Powered Neutron Stars and the Mass–Radius Relation	294
	Problems	303
7	Rapidly Rotating Neutron Stars	307
7.1	Spacetime of Stationary and Axisymmetric Rotating Bodies	308
7.1.1	Physical Interpretation of the Metric	309
7.1.2	Geodetic and Lense–Thirring Precession	312
7.1.3	On General 3+1 Split of Spacetime	315
7.2	Einstein’s Field Equations for Rotating Objects	317
7.2.1	Ricci Tensors of Time-Slices	318
7.2.2	Extrinsic Curvature and 4D Ricci Tensors	319
7.2.3	3+1 Split of Einstein’s Equations	320
7.3	Stellar Structure Equations in Isotropic Gauge	321
7.3.1	The Isotropic Gauge	321
7.3.2	Structure Equations for Rotating Stars	322
7.3.3	Mechanical Equilibrium and Effective Potential	324
7.3.4	Stellar Parameters	326
7.4	The Slow-Rotation Approximation	332
7.5	Numerical Integration of the Stellar Structure Equations	335
7.5.1	Comparison of Numerical Codes	337
7.5.2	Properties of Rotating Equilibrium Stellar Structures	338
7.6	Towards Analytical Vacuum Solutions for Rotating Neutron Stars	342
7.6.1	Weyl–Papapetrou Form	342
7.6.2	Ernst Equations	343
7.6.3	Manko’s Solution	345
7.7	On Oscillation and Formation of Rotating Neutron Stars	350
	Problems	353
8	Black Holes	355
8.1	The Schwarzschild Black Hole	355
8.1.1	Tortoise Coordinates and Null Cones	356
8.1.2	Roads towards Black Hole Formation	358
8.1.3	The Kruskal Extension	359
8.1.4	Penrose Diagram – the Conformal Structure of Infinity	363
8.2	Geodetic Motions in Schwarzschild Spacetime	369
8.2.1	A Lagrangian	369
8.2.2	The Effective Potential for Equatorial Motion	371

8.2.3	Orbital Equation and Bound Orbits in Schwarzschild Spacetime	373
8.3	The Kerr Black Hole	378
8.3.1	Kerr Black Hole in Boyer–Lindquist Coordinates	379
8.3.2	A Short Derivation of the Kerr Solution	379
8.3.3	The Weyl–Papapetrou Form of the Kerr Metric	384
8.3.4	Uniqueness of the Kerr Solution	385
8.3.5	Global Properties of the Kerr Metric	386
8.3.6	On the Conformal Structure of the Kerr Solution	393
8.3.7	Ernst’s Equations for the Kerr Geometry	394
8.3.8	The Kerr–Schild Metric and Two-Black-Hole States	395
8.4	Rotational Energy and the Four Laws of Black Hole Evolution	399
8.4.1	Surface Gravity and Angular Velocity of the Horizon	400
8.4.2	First Law of Black Hole Dynamics	402
8.4.3	Rotational Energy of Astrophysical Black Holes	405
8.4.4	On the Second and Third Laws of Black Hole Dynamics	406
8.5	Time Evolution of Black Holes	408
8.5.1	Quasistationary Evolution of Accreting Black Holes	408
8.5.2	Merging of Black Holes	411
8.6	Geodesics in the Kerr Geometry	412
8.6.1	Direct Integration of Geodesics Equations	414
8.6.2	Geodesics in the Equatorial Plane	416
8.6.3	Geodesics Including Lateral Motion	424
8.6.4	Null Geodesics and Ray-Tracing in Kerr Geometry	431
8.7	Dark Energy Stars	442
8.7.1	Why Dark energy Stars?	442
8.7.2	Structure of Gravastars	443
8.7.3	The Necessity of an Anisotropic Crust	445
	Problems	446
9	Astrophysical Black Holes	449
9.1	Classes of Astrophysical Black Holes	450
9.2	Measuring Black Hole Masses	451
9.2.1	BHs in X-Ray Binaries	451
9.2.2	Intermediate-Mass Black Holes	456
9.2.3	Supermassive Black Holes in Nearby Galaxies	456
9.2.4	Black Holes in Quasars	468
9.3	Estimating Black Hole Spin	470
9.3.1	Black Hole Spin and Radio Galaxies	471
9.3.2	Spectral Fitting of Accretion Disks	471
9.3.3	Relativistic Iron Lines	472
9.3.4	Quasiperiodic Oscillations	472
9.4	Black Holes and Galaxy Formation	472
9.5	Black Hole Magnetospheres	473
9.5.1	The 3+1 Formalism for Maxwell’s Equations	474

9.5.2	Plasma Equations in the 3+1 Split	478
9.5.3	Time Evolution of Magnetic and Current Flux in Turbulent Disks	480
9.5.4	Stationary Magnetospheres on Kerr Black Holes	486
9.5.5	Relaxation of Black Hole Magnetospheres and the Blandford–Znajek Process	499
9.6	Magnetic Spin-Down of Rotating Black Holes	509
	Problems	511
10	Physics of Accretion Flows around Compact Objects	513
10.1	Angular Momentum Transport	514
10.2	Magnetohydrodynamics for Accretion Disks	517
10.2.1	Equations of Magnetohydrodynamics	517
10.2.2	Time and Space Discretization	523
10.2.3	MRI Driven Turbulence in Disks	525
10.2.4	Two-Temperature Plasmas and Radiation Pressure in Accretion Disks	533
10.3	States of Turbulent Accretion Disks	537
10.3.1	Turbulent Angular Momentum Transport in Accretion Disks	538
10.3.2	Truncated Accretion and Standard Disk Models in 1D	540
10.3.3	Standard Thin Disk Solutions (SSD)	545
10.3.4	Advection-Dominated Flows (ADAF)	551
10.3.5	Super-Eddington Accretion	552
10.3.6	Unified Models of Disk Accretion	553
10.3.7	Fundamental Time-Scales for Accreting Black Holes	555
10.4	Relativistic MHD – Turbulent Accretion onto Black Holes	558
10.4.1	From SRMHD to GRMHD	558
10.4.2	The Equations for GRMHD	559
10.4.3	Nonradiative Accretion onto Rotating Black Holes	563
10.5	Jets and the Ergosphere	565
10.5.1	Jets as Outflows from the Ergospheric Region	566
10.5.2	From the Ergosphere to the Cluster Gas	572
	Problems	575
11	Epilogue and Future Prospects	579
	Astrophysical Constants and Symbols	587
	SLy4 Equation of State for Neutron Star Matter	591
	3+1 Split of Spacetime Curvature	595
C.1	Gauss Decomposition	595
C.2	Codazzi–Mainardi Equations	596

3+1 Split of Rotating Neutron Star Geometry	599
D.1 The 3+1 Split of the Connection	599
D.2 The Curvature of Time Slices	601
Equations of GRMHD	605
E.1 Electromagnetic Fields	605
E.2 Conservative Formulation of GRMHD	607
E.3 Numerical Schemes	609
Solutions	613
Glossary	641
References	657
Index	675

List of Figures

1.1	Fe K Line of the Seyfert galaxy MCG–6–30–15	3
1.2	Compact objects in binary systems	6
1.3	Endpoints of stellar evolution	7
1.4	Stellar evolution	8
1.5	Accretion disk in Cyg X–1	10
1.6	Spectral states for Cyg X–1	10
1.7	A radio pulsar	12
1.8	Pulsar nebula around the Crab pulsar	12
1.9	Masses of compact objects	17
1.10	AGN paradigm	18
1.11	Profiles for gamma-ray bursts	20
1.12	Decay of the GRB 970508 flux in X-rays	21
1.13	Light-curves of the short burst GRB060313	22
1.14	The Swift satellite	23
1.15	Distribution of GRBs on the sky	24
1.16	A GRB model	24
2.1	WEP Experiments	32
2.2	Meaning of curvature	57
2.3	Gravity wave experiments	75
2.4	Modes of gravitational waves	86
2.5	GEO600 sensitivity achieved in 2006	87
2.6	Three phases of black hole merger	90
2.7	Slicing of spacetime	94
3.1	Geodesic spray	117
4.1	Metric functions of static stars	132
5.1	The binary system Sirius A and B	138
5.2	The orbit of the binary system Sirius A and B	139
5.3	Spectra of DA white dwarfs	140
5.4	Classification of white dwarfs	141
5.5	White dwarfs found in SDSS	142
5.6	Mass distribution of white dwarfs in SDSS	143

5.7	White dwarfs in the Globular Cluster M4	144
5.8	Magnetism of white dwarfs	144
5.9	Periods in magnetic white dwarfs	145
5.10	White dwarf tracks in the HR diagram	146
5.11	Spectral energy distribution of LHS 1126	147
5.12	WD luminosity function	148
5.13	GAIA WD luminosity function	150
5.14	Internal structure of white dwarfs	152
5.15	White dwarf masses	163
5.16	Mass-radius relations for WDs	163
5.17	White dwarf masses and radii	165
5.18	White dwarf masses and radii for carbon cores	166
5.19	White dwarf masses and radii for iron cores	167
5.20	Core temperature vs. luminosity	176
5.21	Specific heat capacity of diamond	178
5.22	Cooling sequences for white dwarfs	179
5.23	Bolometric magnitudes as a function of time	180
5.24	Structure of a cataclysmic system	181
5.25	Magnetic CVs	184
6.1	Interior of a neutron star	189
6.2	EoS for cold nuclear matter	193
6.3	Empirical EoS for neutron matter	197
6.4	Nuclear EoS compared to empirical ones	200
6.5	Classical neutron stars: EoS	201
6.6	Classical neutron stars: APR density profiles	202
6.7	Classical neutron stars: density profiles	202
6.8	Classical neutron stars: mass vs. central density	203
6.9	Classical neutron stars: mass-radius relation	204
6.10	Classical neutron stars: surface redshift	205
6.11	Classical neutron stars: moment of inertia	206
6.12	EoS for neutron matter	216
6.13	Analytic fits to neutron star EoS	218
6.14	Density as a function of number density	220
6.15	Weak flavor changing reaction	221
6.16	Particles in hadronic stars	222
6.17	Hyperon star radii as a function of mass	223
6.18	Phase diagram for quark matter	224
6.19	Phase diagram for quark matter	234
6.20	Quark cores in neutron stars: EoS	236
6.21	Quark cores in neutron stars: mass vs. radius	238
6.22	Maximum mass of quark stars	238
6.23	The observed X-ray spectrum of EXO 0748–676	240
6.24	Mass-radius relation for strange quark stars	241
6.25	Structure of neutron stars for high masses	243

6.26	Orbital decay of the binary system PSR 1913+16	254
6.27	Geometry and orbital elements	256
6.28	Timing residuals for PSR B1534+12	261
6.29	Modulation of pulsar arrival times for PSR 1913+16.....	262
6.30	Mass diagram for the binary pulsar system B 1534+12	265
6.31	Observed neutron star masses	266
6.32	Mass diagram for the binary pulsar system PSR 0737–3039A+B ..	268
6.33	Shapiro time delay	269
6.34	X-ray spectrum of the neutron star RX J1856–3754	274
6.35	Radius constraints on the mass–radius relation.....	276
6.36	Specific heat in a minimal model of neutron stars	279
6.37	NS cooling curves for APR EoS	280
6.38	NS cooling curves for minimal model	282
6.39	Surface temperature as function of core temperature	283
6.40	NS cooling curves for hybrid stars	283
6.41	Magnetosphere of rotating neutron stars	285
6.42	Pulsar diagram for the ATNF pulsar catalogue	286
6.43	Pulsar period histogram for the ATNF pulsar catalogue	286
6.44	Global spectra of seven radio pulsars	288
6.45	Observed X-ray spectrum of PSR B0656	289
6.46	Pulsars expected to be seen with GLAST	292
6.47	Observed radio and X-ray light-curves	293
6.48	Be binary system	295
6.49	X-ray light-curve and QPOs of 4U1702–429	299
6.50	Power spectrum of Sco X-1	300
6.51	Plot of upper kHz vs. lower kHz QPO	301
7.1	The satellite Gravity Probe B	313
7.2	Geodetic precession and Lense–Thirring effect	314
7.3	NS frame-dragging potential	339
7.4	NS gravitational mass vs. central density	339
7.5	Rotating NS gravitational mass vs. radius	340
7.6	NS sequences of constant rest mass	341
7.7	Existence of Manko solutions	349
7.8	Rotational frequency for Manko solutions	349
8.1	Finkelstein diagram of collapsing star	357
8.2	Kruskal diagram	361
8.3	Minowski embedding into Einstein Universe	365
8.4	Penrose diagram for Minkowski	365
8.5	Penrose diagram for Schwarzschild spacetime	367
8.6	Effective potential	372
8.7	Orbits of the first kind near a black hole	377
8.8	Schwarzschild vs. Kerr black hole	387
8.9	Ergosphere of black holes	390

XXII List of Figures

8.10	Frame-dragging of black holes	391
8.11	Characteristic Kerr radii	392
8.12	Conformal structure of Kerr	394
8.13	Black hole collision	413
8.14	Black hole impact	414
8.15	Black hole impact 2	415
8.16	Black hole impact 3	415
8.17	Angular velocity of quasi-Keplerian orbits: slow rotation	423
8.18	Angular velocity of quasi-Keplerina orbits: rapid rotation	423
8.19	Ray-tracing in the Kerr geometry	432
8.20	Ray-tracing in Kerr geometry	433
8.21	Disk redshift distribution	438
8.22	Profile of the broad iron line	440
8.23	Relativistic line profiles	441
8.24	Relativistic iron line profiles vs. a	441
8.25	Schematic pressure profile of a gravastar	444
8.26	Volume function of gravastars	446
9.1	X-ray nova outburst GRO J1655–40	450
9.2	Light-curves of the microquasar GRO 1915+105	453
9.3	Microquasars vs. quasars	454
9.4	BH X-ray binary systems	455
9.5	Maser emission in NGC 4258	459
9.6	Galactic center and its black hole	460
9.7	Stellar motion in GC	461
9.8	Central mass distribution in the Galaxy	462
9.9	Spectrum of Sgr A*	463
9.10	Emission from Sgr A*	464
9.11	Correlation between black hole mass and bulge mass	465
9.12	Correlation between black hole mass and velocity dispersion	466
9.13	Correlation between black hole mass and velocity dispersion	467
9.14	Black hole masses for AGNs from reverberation	469
9.15	Quadrupolar field topology around a rapidly rotating black hole	483
9.16	Light cylinder surfaces in Kerr	490
9.17	Field line rotation in Kerr	492
9.18	Axisymmetric magnetosphere of neutron stars	492
9.19	Magnetic flux in the meridional plane of a black hole	493
9.20	Magnetic accretion onto black holes	495
9.21	Magnetic wind solution	497
9.22	Monopole magnetosphere for black holes	498
9.23	Evolution of a split monopole of black holes	503
9.24	Evolution of a Wald magnetosphere	504
9.25	Induction in BH magnetosphere	506
10.1	A quad-tree in 2D adaptive mesh refinement	524

10.2	MRI growth rates	527
10.3	Compact objects in binary star systems	528
10.4	Torus evolution including weak magnetic fields	529
10.5	Accretion torus around black holes	530
10.6	MRI in action	531
10.7	QPOs from magnetorotational turbulence	532
10.8	Photon bubbles in accretion disks	534
10.9	Thin disk solutions	541
10.10	States of accretion disks	545
10.11	Standard disk spectra	549
10.12	Non-LTE disk spectra	550
10.13	ADAF model spectrum	551
10.14	Schematic sketch of the accretion flow	554
10.15	PDSs of the Seyfert galaxy NGC 4051	557
10.16	Break time-scales for BHs	557
10.17	Truncated accretion on black holes	558
10.18	Disk evolution near rotating black holes	564
10.19	Magnetic fields near rotating black holes	565
10.20	Jets from the ergosphere	567
10.21	Jets from ergosphere	568
10.22	M87 in the Virgo cluster	568
10.23	AGN jets	571
10.24	Cygnus A and its cluster gas	573
10.25	The quasar 3C273	574
11.1	Black hole explorers	583
F.1	Energy spectrum of the Crab Nebula	620

1 Compact Objects in Astrophysics

As a class of astronomical objects, compact objects include white dwarfs, neutron stars and black holes. As the endpoint states of stellar evolution, they form today fundamental constituents of galaxies. In the form of supermassive black holes, these objects also live in practically every center of a galaxy. Our Galaxy harbors a black hole of 3.8 million solar masses, but the center of M87 in the Virgo cluster encloses a black hole of three billion solar masses. These supermassive black holes are the most extreme objects found in the Universe.

While neutron stars and stellar mass black holes mainly entered astrophysical research by means of their radio and X-ray emission, white dwarfs had already been detected 100 years ago by their optical emission.

1.1 Why is Newtonian Gravity Obsolete?

The gravitational collapse of normal matter produces some of the most exotic objects in the Universe – neutron stars and black holes. Proving that these objects exist in Nature occupied theoretical and observational astrophysicists for much of the 20th century. Most of the detailed debate centered around understanding the possible final states of massive stars. On his now famous sea voyage from India to England in 1930, Subrahmanyan Chandrasekhar considered the structure of white dwarf stars – compact stellar remnants in which gravitational forces are balanced by electron degeneracy pressure. He realized that, if the white dwarf was sufficiently massive, the degenerate electrons will become relativistic thereby rendering the star susceptible to further gravitational collapse. Although hotly debated by Arthur Eddington, Chandrasekhar correctly deduced that a white dwarf would undergo gravitational collapse if its mass exceeded $M_{\text{Ch}} \simeq 1.4 M_{\odot}$ (where M_{\odot} is the mass of the Sun), a limit now known as the **Chandrasekhar limit**.

Once gravity overwhelms electron degeneracy pressure, neutron degeneracy pressure is the last, best hope for averting total gravitational collapse. Objects in which gravitational forces are balanced by neutron degeneracy pressure are called neutron stars. Although there was initial hope that nuclear forces would always be sufficient to resist gravity, the upper limit to the mass of a neutron star is now believed to be in the range $(1.5\text{--}2.2) M_{\odot}$. Uncertainties arising from the equation of state at supranuclear densities continue to plague our determination of this critical mass, but an absolute upper limit of $3 M_{\odot}$ arises from very general considerations,

i.e. the validity of general relativity and the principle of causality. Above this mass, it is thought that complete gravitational collapse cannot be avoided. In particular, Hawking's singularity theorems show that the formation of a spacetime singularity is unavoidable (irrespective of the mass/energy distribution) once the object is contained within the light-trapping surface. The result is a black hole, i.e. a region of spacetime bounded by an event horizon and, at its heart, possessing a spacetime singularity.

While the above considerations now have a firm theoretical base, observational astrophysics was, and continues to be, critically important in guiding our understanding of such extreme objects. In the case of both neutron stars and black holes, the very existence of these objects was only widely accepted when compelling observational evidence was forthcoming. For neutron stars, the pivotal observation was the discovery of radio pulsars by Jocelyn Bell and Anthony Hewish via radio observations taken from Cambridge. Black holes gained wide acceptance after it was demonstrated that the X-ray emitting compact object in the binary star system **Cygnus X-1** did, in fact, possess a mass in excess of the maximum possible neutron star mass. This made it the first of the so-called Galactic Black Hole Candidates (GBHCs), a class that has now grown to include some two dozen objects.

We now know of another class of black holes – the **supermassive black holes**, with masses in the range of $(10^6\text{--}10^{10}) M_\odot$, that reside at the dynamical centers of most, if not all, galaxies. Today, by far the strongest case for a supermassive black hole can be made for our own Galaxy. Modern high-resolution, infrared imaging reveals that the stars in the central-most regions of our Galaxy are orbiting an unseen mass of three million solar masses. Furthermore, studies of the orbital dynamics (which now include measured accelerations as well as velocities) constrain the central mass to be extremely compact. According to conventional physics, the only long-lived object with these properties is a supermassive black hole. Alternatives, such as a compact cluster of neutron stars, would suffer a dynamical collapse on much shorter time-scales.

Having established beyond reasonable doubt that black holes exist, it is obviously interesting to perform detailed observational studies of them. The regions in the immediate vicinity of a black hole bear witness to complex interactions between matter moving at relativistic velocities, electromagnetic fields, and the black hole spacetime itself. Given that the apparent angular scales of even the biggest black hole event horizons are $\simeq 10^{-6}$ arcsec, direct imaging studies of these regions will not be possible for many years. In the meantime, we must study these regions using more indirect methods, chief among which are spectroscopic methods.

As we will detail in this book, Nature has provided us with a well-understood and extremely useful spectral diagnostic of matter in the near vicinity of astrophysical black holes. In essence, relatively cold matter in the near vicinity of an astrophysical black hole will inevitably find itself irradiated by a spectrum of hard X-rays. The result can be a spectrum of fluorescent emission lines, the most prominent being the $K\alpha$ line of iron at an energy of 6.4 keV (depending upon the ionization state of the iron). Ever since the launch of the Advanced Satellite for Cosmology and Astrophysics (ASCA) in February 1993, X-ray astrophysicists have had the capability

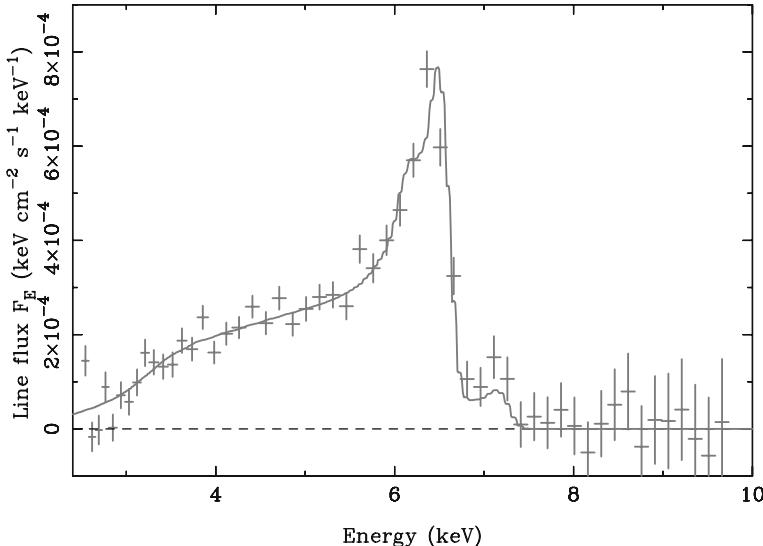


Fig. 1.1. Fe K Line of the Seyfert galaxy MCG–6–30–15 as observed with XMM–Newton. The line is intrinsically narrow with a rest-frame energy of 6.4 keV. The line originates from material that is just a few gravitational radii from the black hole, its profile is shaped by Doppler shifts and gravitational redshift effects. Figure adapted from [150]

to identify this emission line and measure its spectral profile. Figure 1.1 shows the iron line in the X-ray emissions originating near the supermassive black hole in the galaxy MCG–6–30–15. Bearing in mind that the line is intrinsically narrow with a rest-frame energy of 6.4 keV, it can be seen that the line has been dramatically broadened and skewed to low energies. It is now widely accepted that the line originates from material that is just a few gravitational radii from the black hole, and possesses a profile that is shaped by (relativistic) Doppler shifts and gravitational redshift effects. Investigating these spectral features in X-ray luminous black hole systems has given us the clearest window to date on the physics that occurs in the immediate vicinity of astrophysical black holes.

1.2 Einstein was Skeptical about the Existence of Black Holes

A black hole is a region of space whose attractive gravitational force is so intense that no matter, light, or communication of any kind can escape. A black hole would thus appear black from the outside. However, gas around a black hole can be very bright. It is believed that black holes form from the collapse of stars. As long as they are emitting heat and light into space, stars are able to support themselves against their own inward gravity with the outward pressure generated by heat from nuclear reactions in their deep interiors. Every star, however, must eventually exhaust its nuclear fuel. When it does so, its unbalanced self gravitational attraction causes it

to collapse. According to theory, if a burned-out star has a mass larger than about twice the mass of our Sun (as a protoneutron star), no amount of additional pressure can stave off total gravitational collapse. The star collapses to form a black hole. For a nonrotating collapsed star, the size of the resulting black hole is proportional to the mass of the parent star; a black hole with a mass three times that of our Sun would have a diameter of about 20 km. The possibility that stars could collapse to form black holes was first theoretically discovered in 1939 by J. Robert Oppenheimer and H. Snyder [318], who were manipulating the equations of Einstein's general relativity. The first black hole believed to be discovered in the physical world, as opposed to the mathematical world of pencil and paper, was Cygnus X-1, about 7000 lightyears from Earth. Cygnus X-1 was found in 1970. Since then, a few dozens excellent black hole candidates have been identified. Many astronomers and astrophysicists believe that massive black holes, with sizes up to 10 billion times that of our Sun, inhabit the centers of energetic galaxies and quasars and are responsible for their enormous energy release. Ironically, Einstein himself did not believe in the existence of black holes, even though they were predicted by his theory.

1.3 Subrahmanyan Chandrasekhar and Compact Objects

Subrahmanyan Chandrasekhar was born in Lahore (then in British India) and studied physics at the Presidency College, Madras. In 1930, he became a research student of R.H. Fowler at Cambridge University and earned his PhD in 1933. He developed the theory of white dwarf stars, showing that quantum mechanical degeneracy pressure cannot stabilize a massive star. He showed that a star of a mass greater than 1.4 times that of the Sun (now known as the Chandrasekhar limit) had to end its life by collapsing into an object of enormous density such as a black hole. In 1937, he joined the University of Chicago and the Yerkes Observatory. He investigated and wrote important books on stellar structure and evolution, dynamical properties of star clusters and galaxies, radiative transfer of energy, hydrodynamic and hydromagnetic stability, the stability of ellipsoidal figures of equilibrium, and the mathematical theory of black holes. He also worked in relativistic astrophysics, and his last book was *Newton's Principia for the Common Reader*. In 1952, he received the Catherine Wolfe Bruce gold medal, for lifetime contributions to astronomy. He was awarded the Royal Medal of the Royal Society in 1962, and he edited the *Astrophysical Journal* for nearly 20 years. Chandrasekhar shared the 1983 Nobel Prize in physics with W.A. Fowler for his studies of the physical processes of importance to the structure and evolution of stars.

Chandrasekhar left Bombay on a boat on 31st July 1930. On the voyage, after overcoming his seasickness, he remembered Fowler's paper and decided to combine it with his knowledge of special relativity theory. To his great surprise, he found that this combination predicted that white dwarfs could only exist up to a certain limiting mass which depended chiefly on fundamental constants such as h , G and the mass of the hydrogen atom; the mass was about 1.45 times the mass of the Sun. England's two leading astrophysicists, Eddington and Milne, could not believe this

result, and neither of them would recommend Chandra's paper for publication by the Royal Society. So Chandra sent it to the Astrophysical Journal in America, which published it in March 1931.

Of course, Eddington was wrong. But his resistance to Chandra's mass limit was understandable: his life's work had been to show that every star, whatever its mass, had a stable configuration. It was generally (and correctly) believed that white dwarfs were the end stage of stellar evolution, after their energy source was exhausted. Why should there be a limit to the mass of a star in its old age? Chandra appealed to physicists he knew – Rosenfeld, Bohr, Pauli. Unanimously, they decided that there was no flaw in his argument. But it took decades before the Chandrasekhar limit was accepted by the astrophysics community.

The paradox that *normal stars* can exist with any mass whereas white dwarfs can only exist up to 1.45 solar masses is now understood. Stars, in their evolution, go through a giant stage in which their radius may be hundreds of times larger than originally. In this stage, the atoms at the surface are not strongly held by gravity, while there is strong radiation pressure from the inside. Some atoms, especially hydrogen, are blown off and the star gradually loses mass. Theory shows that stars up to eight solar masses lose mass in this manner, ending up below the Chandrasekhar limit. None of this was known in 1935.

The limit also affects stars heavier than eight solar masses. Matter in the central core of stars evolves to iron by successive nuclear reactions. At this point, no further nuclear energy can be obtained, just as in white dwarfs. When the iron core grows to the Chandrasekhar mass, it collapses by gravitation into a neutron star, and the rest of the star is expelled, giving a type II supernova. Some white dwarfs accrete matter from the outside, and when their mass has grown to the Chandrasekhar limit, they also become supernovae, in this case type Ia. Chandra's theory is basic to much modern astrophysics.

1.4 Classes of Compact Objects

The study of compact stars begins with the discovery of white dwarfs and the successful description of their properties by the Fermi–Dirac statistics, assuming that they are held up against gravitational collapse by the degeneracy pressure of the electrons, an idea first proposed by Fowler in 1926 [160]. A maximum mass for white dwarfs was found to exist in 1930 by the seminal work of Chandrasekhar due to relativistic effects [113]. In 1932 Chadwick discovered the neutron. Immediately, the ideas formulated by Fowler for the electrons were generalized to neutrons. The existence of a new class of compact stars, with a large core of degenerate neutrons, was predicted – the neutron stars (NS). The first NS model calculations were achieved by Oppenheimer and Volkoff [317] and Tolman [394] in 1939, describing the matter in such a star as an ideal degenerate neutron gas. Their calculations also showed the existence of a maximum mass, like in the case of white dwarfs, above which the star is not stable and collapses into a black hole. They found a maximum stable mass of $0.75 M_{\odot}$ [317]. Only nearly 30 years later, in 1967, was the first neutron

star observed – in fact, a strange object pulsating in the radio range (radio pulsar), which was however quickly identified as a fast rotating neutron star. Already in 1964, black holes have been proposed as the ultimate energy source for quasars. In the meantime, the existence of black holes has been established in a huge mass range, from about three solar masses to 10 billion solar masses in the centers of huge elliptical galaxies.

In 1974, the pulsar PSR 1913+16 was observed for the first time in a binary system by Hulse and Taylor. This allowed a precise measurement of its mass which was found to be $1.44 M_{\odot}$. Hence, this mass measurement ruled out the simple picture of an ideal gas of neutrons for the interior of this star. It shows that the interactions between the nucleons must be taken into account.

Shortly after the introduction of the quark model for nucleons, theoreticians speculated about the possible existence of quark matter inside neutron stars. Gerlach demonstrated in his PhD thesis with Wheeler in 1968 [170] that a third family of compact stars could exist in Nature, besides white dwarfs and neutron stars. He derived general conditions on the equation of state for such a new form of stars to exist, in particular that a strong softening in the equation of state, like in phase transitions, has to occur in neutron stars. Some astrophysicists even argue that the very ground state of matter is in fact strange quark matter (composed of u, d and s quarks). Such objects have now been studied since the mid-1980s and are referred to as **strange stars** [27, 186].

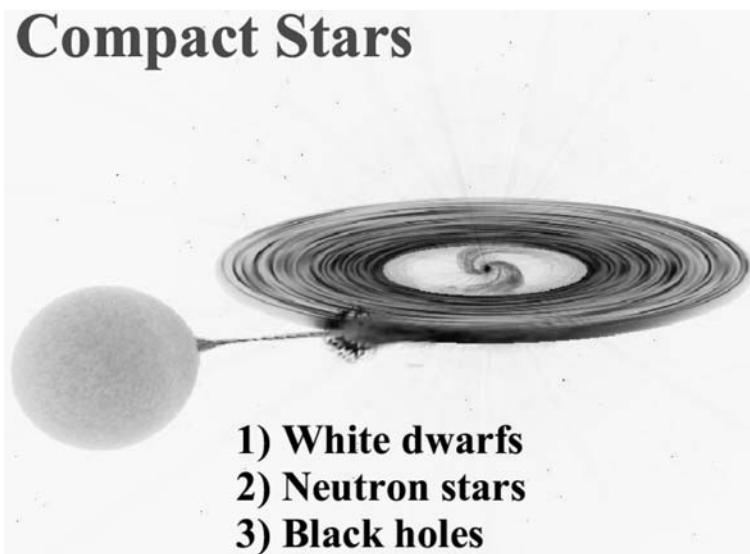


Fig. 1.2. Compact objects in binary systems. Compact stars are formed in stellar evolution and often live as companion in a binary system, surrounded by gas rings formed by mass overflow from its companion star

A study of compact objects – white dwarfs, neutron stars, and black holes – begins when normal stellar evolution ends. All these objects differ from normal stars in at least two aspects:

- They are not burning nuclear fuel, and they cannot support themselves against gravitational collapse by means of thermal pressure. Instead white dwarfs are supported by the pressure of the degenerate electrons, and neutron stars are largely supported by the pressure of the degenerate neutrons and quarks. Only black

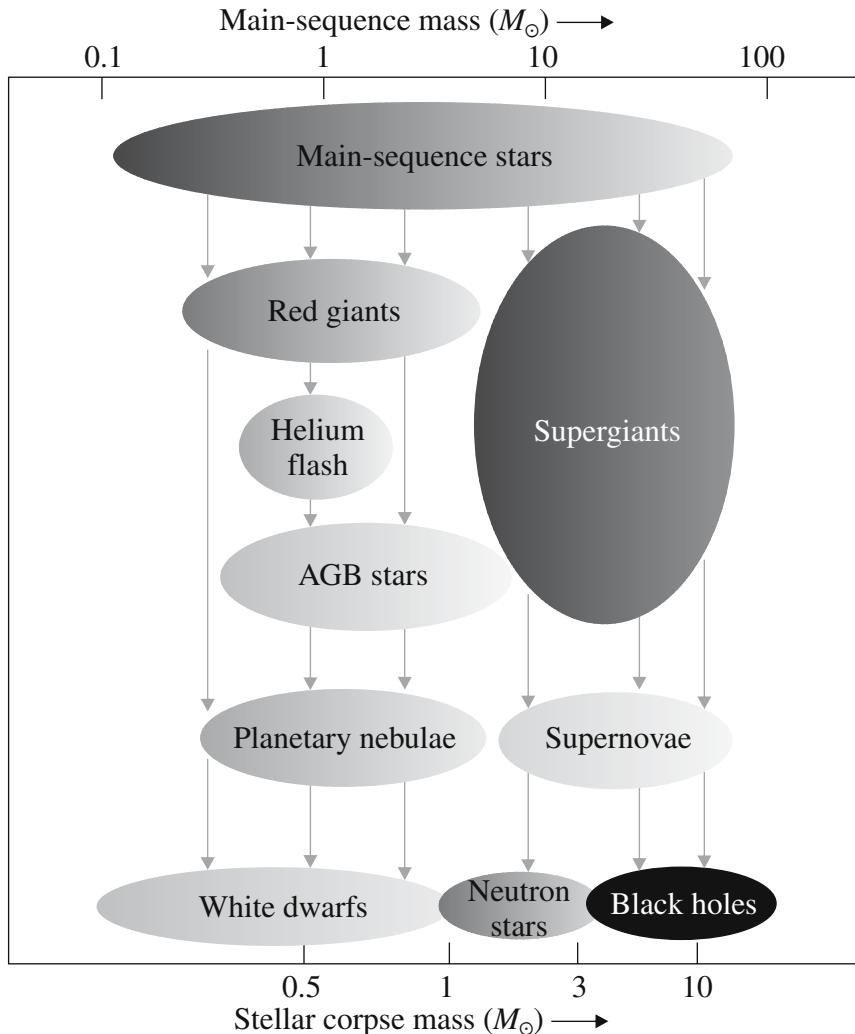


Fig. 1.3. Compact stars are the result of the endpoint in stellar evolution shown as a function of the initial mass

holes represent completely collapsed stars, assembled by mere self-gravitating forces. These objects can be considered as a kind of soliton solution of Einstein's equations.

- The second characteristic property of compact stars is their compact size. They are much smaller than normal stars and therefore have much stronger surface gravitational fields.
- Often compact objects carry strong magnetic fields, much stronger than found in normal stars.

1.4.1 White Dwarfs and Neutron Stars

White dwarfs are stars of about one solar mass with a characteristic radius of 5000 km, corresponding to a mean density of 10^6 g cm^{-3} . They are no longer burning nuclear fuel, but are steadily cooling away their internal heat. In 1926, only three white dwarfs were firmly detected. In that year, Dirac formulated the Fermi–Dirac statistics, which was used by Fowler [160] in the same year, in a pioneering paper on compact stars – to explain the puzzling nature of white dwarf stars. He identified the

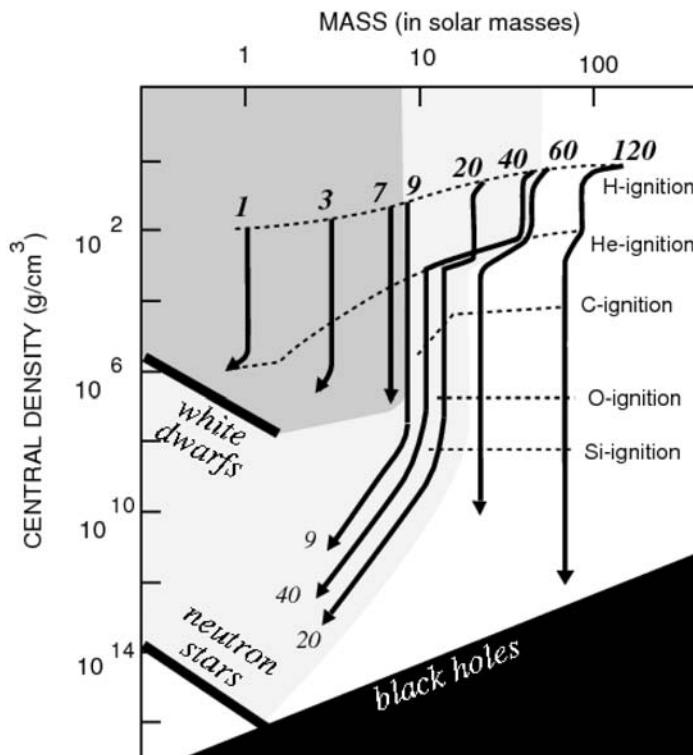


Fig. 1.4. Stars evolve towards different end states: white dwarfs, neutron stars and black holes. Figure adapted from [260]

pressure holding up the stars from gravitational collapse with the electron degeneracy pressure.

Actual models of white dwarf stars, taking into account the special relativistic effects in the degenerate electron equation of state were then constructed in 1930 by Chandrasekhar [113]. He made the fundamental discovery of a maximum mass of $1.4 M_{\odot}$ for white dwarfs – the exact value somewhat depends on the chemical composition.

The prediction of the existence of neutron stars as a possible endpoint of stellar evolution was independent of observations. Following the discovery of the neutron by Chadwick, it was realized by many people that at very high densities electrons would react with protons to form neutrons via inverse beta decay. Neutron stars had been found at the end of the 1960s as radio pulsars and in the beginning of the 1970s as X-ray stars. A firm upper limit for the mass of neutron stars was then seen as evidence for the existence of even more exotic objects – black holes. At the time of the discovery of Cyg X-1 by Uhuru (1970) the value of this upper limit was, however, the subject of great debate.

1.4.2 Compact X-Ray Sources

A new era in astronomy was opened up in the 1960s by means of the launch of various rockets (Giacconi 1962). They discovered Sco X-1 in the energy band of 1–10 keV. At the end of one decade, about 20 X-ray sources had been identified. One of the strongest sources, Cyg X-1, was also found to vary in time. Already at that time, gas accretion in a close binary system was seen to be the source of this X-ray emission. But, for example, Prendergast and Burbidge [332] argued that gas flowing onto a compact star in a binary system would have too much angular momentum to flow radially inwards. They suggested that the gas would form a disk around the compact star, with approximately Keplerian angular momentum. There should exist a small inward drift velocity. The notion of an accretion disk was born (Fig. 1.5). A comprehensive and up-to-date survey on compact stellar X-ray sources written by leading experts in the field can be found in the book [6]. This book covers the details of recent developments in X-ray astronomy and multiwavelength observations, as well as some theoretical issues for these objects.

Modes of Accretion

Over the years, Cyg X-1 has been found to show two pronounced X-ray states (Fig. 1.6):

- a soft or high state: Here a pronounced black-body (BB) spectrum is visible with a temperature of about 1 keV, and the luminosity is high;
- a hard or low state: the BB disappears and the X-rays are emitted in the hard X-ray region up to 150 keV.

The existence of these two states is generally interpreted as evidence for two different modes of accretion. The high energies of the photons is seen as evidence for the

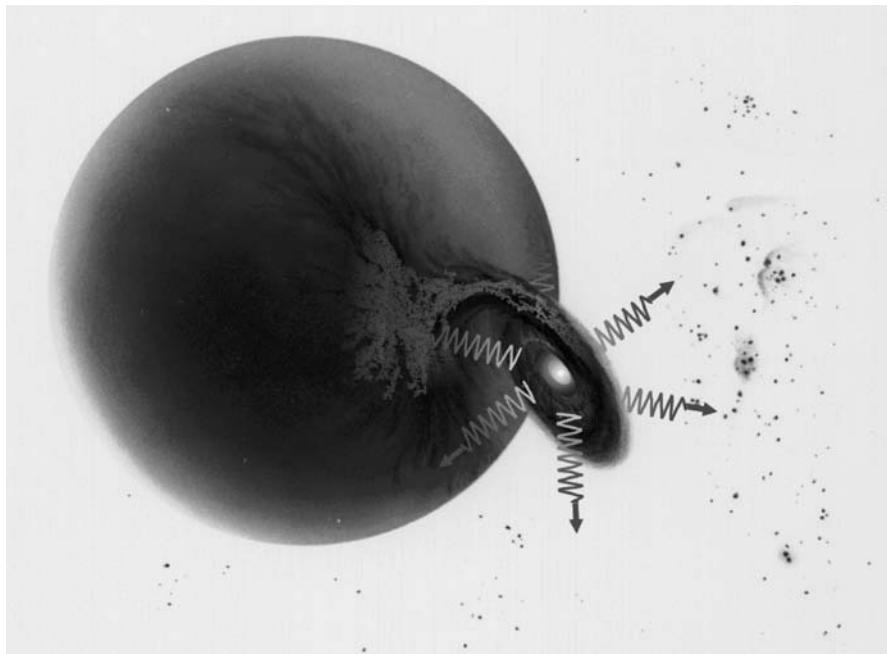


Fig. 1.5. Formation of an accretion disk in a binary system resulting in X-ray emission (wavy lines) and the launch of jets perpendicular to the disk. The companion star of Cygnus X-1 is a blue supergiant

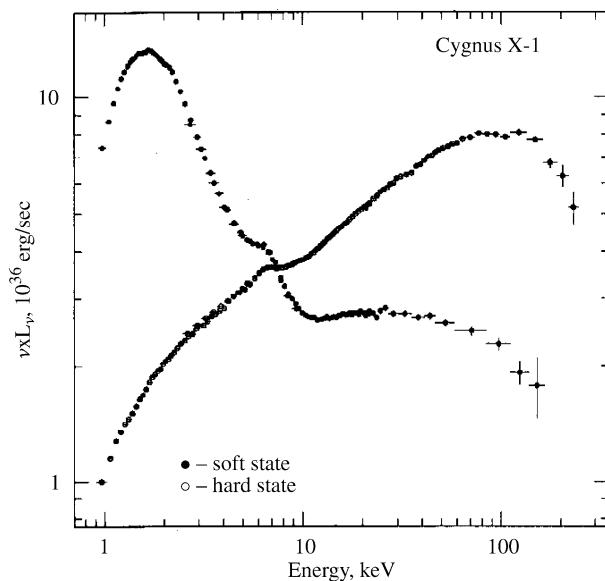


Fig. 1.6. Spectral states for Cyg X-1. The energy distribution νL_ν is shown as a function of the X-ray energy in keV for two different epochs, corresponding to a soft state and a hard state. At 6.4 keV an iron line is visible

existence of a hot plasma with electron temperatures of 10^9 K in the neighborhood of the accretion disk. Soft photons from the optically thick disk would then be up-scattered by Compton processes. Repeated Compton scattering can explain the power-law form of the observed spectra in the hard state. The very location of the hot plasma is still under debate. But already in 1977 Liang and Price introduced the concept of a hot dissipative corona above the accretion disk following the example of the solar corona (Liang and Price [254]). The energy could be dissipated by MHD waves, or jets. In recent years, this transition between high state and low state is seen as a transition from an optically thick disk flow to an optically thin disk flow. This latter mode of accretion always exists within the marginal stable orbit, since matter has to flow with the speed of light through the horizon and has to be supersonic when entering the horizon (for causality reasons). In this way, the inner accretion onto a black hole has to be hot (i.e. high sound speed near the speed of light). When the accretion rate is high, then the soft flux is dominant and cooling of the corona is efficient. On the other hand, when the accretion rate is very low, the inner disk is probably in a very hot state cooling by Comptonization of soft photons from the outer disk; this would correspond to the low state (hard) spectrum.

The existence of these luminosity states is not only generic for black hole systems. Also neutron stars in low-mass X-ray binary systems (LMBXBs) are found to dispose such luminosity states. The difference is the missing hard tail in the soft state in the case of neutron stars, while the soft spectrum of Cyg X-1 has a pronounced hard excess extending to at least a few hundred keV and probably into the MeV region.

1.4.3 Radio Pulsars

Pulsars are the lighthouses of the Galaxy – rapidly spinning neutron stars whose strong magnetic fields produce conical beams of electromagnetic radiation that sweep past the Earth with each rotation of the star, producing the eponymous pulses that are observed primarily at radio wavelengths (Fig. 1.7). Pulsars were discovered, albeit accidentally, by Jocelyn Bell at Cambridge in 1967. The apparently sporadic bursts of radio emission appeared during the course of a survey to investigate the effects of interplanetary scintillation of radio sources. Working as a graduate student in a team lead by Anthony Hewish, Bell soon realized that the emission always occurred at the same position in the celestial sphere indicating that the source was not of terrestrial origin. Subsequent observations with greater time resolution showed the emission to be a train of pulses with a precise repetition period of 1337 ms. The Cambridge team published their discovery the following and, soon afterwards, announced the discovery of three more pulsars found from subsequent inspection of the remaining survey data.

Hewish was awarded the 1974 Nobel Prize in physics for his “decisive role in the discovery of pulsars” and his pioneering work in radio astronomy. Bell’s key role in the discovery has been widely recognized: among other awards, she has received the Michelson Medal of the Franklin Institute in Philadelphia (jointly with Hewish), the Tinsley Prize of the American Astronomical Society, and the Herschel Medal from the Royal Astronomical Society.

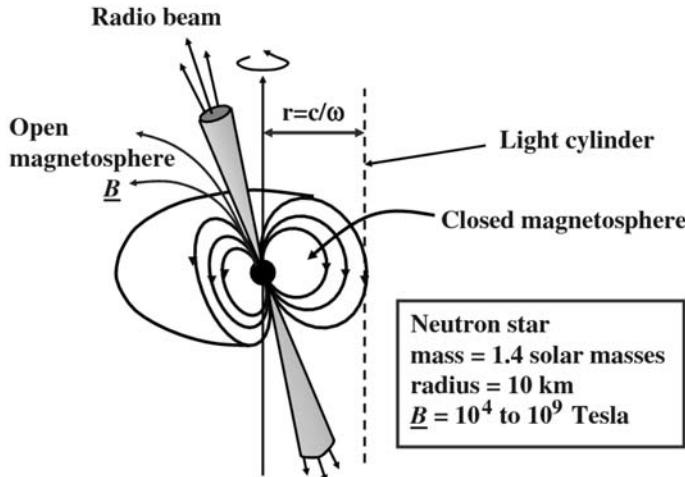
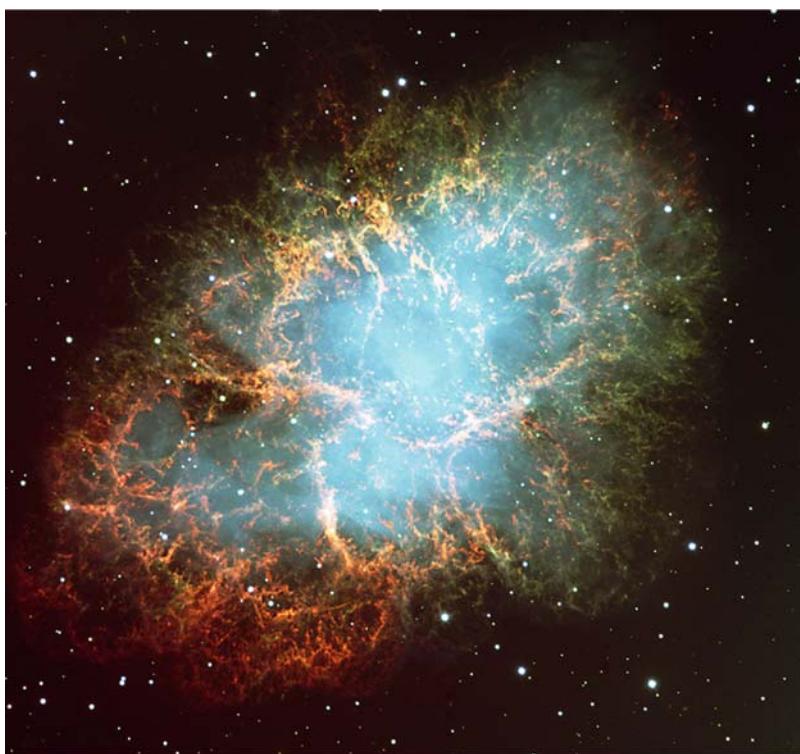


Fig. 1.7. A radio pulsar is a rotating magnetized neutron star with a magnetic dipole moment inclined with respect to the rotational axis. The magnetosphere is closed within the light cylinder $R_L = c/\omega$, where $\omega = 2\pi/P$ is the angular velocity of the neutron star and c the speed of light. Plasma flows away along the open magnetic field lines and emits thereby radio waves. The surface field strengths are between 10^4 and 10^9 tesla (10^8 and 10^{13} gauss)

Observational progress proceeded quickly, with many astronomers throughout the world using the largest radio telescopes to search for more pulsars so that, by the end of 1968, about 20 pulsars were known. By this time, the most plausible explanation for the phenomenon put forward was that pulsars are rapidly rotating, highly magnetized neutron stars radiating energy out of their magnetic poles. In this “lighthouse” model, the observed pulses are produced as the magnetic axis crosses our line-of-sight once per rotation.

Prior to the discovery, neutron stars were a purely theoretical concept – first proposed by Walter Baade and Fritz Zwicky to be the collapsed remains of a massive star after it has exploded as a supernova. The Baade–Zwicky prediction that neutron stars would be associated with supernova remnants was dramatically confirmed with the discovery of a short period (33 ms) pulsar in the Crab Nebula in 1968 (Fig. 1.8). The nebula is the remains of a nearby supernova explosion witnessed by

Fig. 1.8. Top: The Crab Nebula, filled with gaseous filaments, is the result of a star that was seen to explode in 1054 AD. Red indicates the electrons are recombining with protons to form neutral hydrogen, while blue indicates synchrotron emission from the inner nebula. The *lower image* shows synchrotron emission of plasma accelerated by tremendous electric voltages created by the central pulsar (red: radio emission, green: visible emission, blue: X-ray emission). The *inner ring*, with prominent knots, of this X-ray nebula is about one lightyear in diameter. The Crab pulsar is the hot spot in the center of the torus-like structure. Image Credit: NASA and The Hubble Heritage Team (STScI/AURA)



Chinese astronomers in 1054 AD. Using the rotating neutron star model, Thomas Gold of Cornell University, USA, was able to show that the Crab pulsar is the dominant energy supply to its surrounding nebula. The connection between pulsars and rotating neutron stars is now universally accepted.

Like other neutron stars, radio pulsars are born in the supernova explosions that accompany the collapse of massive stars. The nascent pulsars are born rotating at up to about one hundred times per second. It is this stored rotational kinetic energy that powers the pulsar, so like a spinning top the pulsar gradually slows down, reaching spin periods of about a second within a few million years. Eventually, within about 100 million years, the pulsar is spinning too slowly to maintain its radio emission, and it fades from view. Some old pulsars that have binary companions can be “recycled,” or spun back up to fast rotation periods by mass transfer from their companions. Because the resulting **millisecond pulsars** have relatively low magnetic field strengths and hence low energy-loss rates, they can continue to spin rapidly for times that are long compared to the age of the galaxy.

In over 35 years since the discovery, pulsars have proved to be exciting objects to study and, presently, over 1500 are known. Most of these are *normal* in the sense that their pulse periods are of order one second and, with few exceptions, are observed to increase secularly at the rate of about one complete period in 1,000,000,000,000,000! This is naturally explained as the gradual spin-down of the neutron star as it radiates energy at the expense of its rotational kinetic energy. A small fraction of the observed sample are the so-called *millisecond pulsars* which have much shorter periods (< 20 ms) and rates of slowdown of typically only one period in 10,000,000,000,000,000, proving to be extremely accurate clocks. In addition, some pulsars are known to be members of binary systems in which the companion is another neutron star, a white dwarf, or even a main sequence star.

Just over 1500 radio pulsars are now known, all in our own Galaxy except for a few pulsars detected in the Magellanic Clouds. They are studied because neutron stars are intrinsically interesting astronomical objects, but also because the study of pulsars is deeply intertwined with many different branches of both astronomy and physics. Pulsars are, for example, very useful astrophysical probes. For example, a sharp radio pulse emitted by a pulsar is delayed and broadened during its propagation through the dispersive, turbulent interstellar medium, in a way that depends on the frequency and polarization of the signal as well as the properties of the medium. Multifrequency studies of pulsar signals have been used to map the distribution and turbulence structure of ionized material in the Galaxy, as well as the average Galactic magnetic field.

Pulsars are also, by virtue of their very regular, clock-like pulses, useful probes of the gravitational environments in which they are found. The Doppler shifts of the signals from pulsars in binaries can be used to study the binary properties, just as spectral lines are used with normal stars. Some very close pulsar binaries have orbits that are substantially deformed from Keplerian ellipses by general relativistic effects; in these systems, very precise tests of “post-Keplerian” gravity theory have been possible.

The first binary pulsar was discovered by Russell Hulse and Joseph Taylor in 1974, during a survey for new pulsars done at the Arecibo Observatory as part of Hulse's PhD thesis work. Follow-up observations showed that the pulsar is in a high eccentricity ($e = 0.6$), short period (7.8 hours) orbit with another star, which is almost certainly a second neutron star. It was immediately realized that the high velocities and strong gravitational fields in this binary make this object an extraordinary laboratory for studying fundamental physics. In 1993, Hulse and Taylor were awarded the Nobel Prize in physics *for the discovery of a new type of pulsar; a discovery that has opened up new possibilities for the study of gravitation.*

The most famous application of pulsar timing techniques has been to tests of experimental gravitation. In most cases, binary orbits are well approximated as Keplerian ellipses. The high velocities ($\simeq 0.001 c$) and strong gravitational fields in some binary pulsar systems cause relativistic deviations from Keplerian motion to be significant. Five relativistic corrections have been measured: the advance of the angle of periastron of the elliptical orbit (as is seen in the orbit of Mercury); the combined effect of the transverse Doppler shift and the changing gravitational redshift as the eccentric orbit carries the pulsars closer and further from its companion; two parameters describing the Shapiro time delay of the pulsar signal as it propagates through the gravitational potential well of the companion; and the decay of the binary orbit due to gravitational radiation back reaction. The measurement of any two of these effects allows the amplitude of the other three to be predicted, making possible very precise tests of general relativity and alternative gravity theories. This was the subject of the 1993 Nobel Prize in physics, discussed above.

Observed pulsar radio luminosities, together with the small source size, imply extraordinarily high brightness temperatures – as high as 10^{31} K. To avoid implausibly high particle energies, coherent radiation processes are invoked. A maser-like mechanism, involving particles bunched in momentum space, is attractive, if only because maser action has been observed elsewhere in astrophysics, but models with coherent emission from bunches of particles have also been widely discussed, with a bunch of N particles localized in physical space radiating power proportional to N^2 . Coherence by bunching is seen in terrestrial lightning flashes. In detail, severe problems remain in understanding pulsar emission by either the maser or bunching models, and no consensus has emerged.

The emission mechanism itself also remains uncertain. Charged particles gyrating around magnetic field lines produce synchrotron radiation. In the strong magnetic fields of the pulsar magnetosphere, a particle will quickly radiate away its components of momentum that are perpendicular to the field lines, so will be confined to the lowest Landau level. Roughly, the charged particle can be pictured as a bead on a wire, along which the bead is free to move. As a particle moves out along a curved field line, it will produce synchrotron-like radiation that is conventionally called “curvature radiation.” Coherent curvature radiation is currently the most widely accepted model for pulsar radio emission, but many other possible models have been discussed, including models based on relativistic plasma instabilities that are variants of the mechanisms proposed for type III solar radio bursts.

The coherent radio emission from pulsars is a nearly insignificant fraction of the total luminosity inferred from their spin-down rates. A handful of radio pulsars, primarily the youngest, also have pulsed emission at optical, X-ray, and gamma-ray energies. The Crab pulsar, for example, has a spectrum that has been measured all the way from 10 MHz radio waves to at least 5 TeV gamma-rays; the emitted power peaks in the hard X-ray band. The pulsar in the Vela supernova remnant has a spectrum that peaks in the gamma-rays, where it is the brightest object in the sky. (These rotation powered high-energy pulsars should not be confused with X-ray binary stars that are accreting material from a companion.) One pulsar is known, Geminga, that appears quite ordinary except that it is seen only at high energies and not in the radio. It is believed that Geminga is a very nearby radio pulsar whose radio beam does not intersect the Earth. As noted above, some models place the generation of the high-energy emission from pulsars in an outer gap, rather than the polar cap region where the radio emission probably originates.

Neutron stars and pulsars also play a crucial role in understanding subnuclear physics. The average density of a neutron star is comparable to the density in heavy atomic nuclei, and the density reaches several times this value at the core. Neutron stars are thus useful laboratories for studying nuclear materials at densities beyond what can be reached in the laboratory. One topic of considerable interest is the maximum neutron star mass, above which the star becomes unstable to collapse to a black hole. Pulsar timing measurements allow, in some cases, very accurate stellar mass estimates. The most massive known neutron star in a radio pulsar binary is 1.44 solar masses, considerably below the 2.2 solar mass limit of a typical nuclear equation of state.

A more direct probe of neutron star interiors comes from study of timing “glitches,” that are observed in some young pulsars. During a glitch event, the pulsar period decreases in a step function, with amplitude from a few parts in 10^9 to a part in 10^6 . At the same time, the spin-down rate increases, and the pulsar recovers over a time period of weeks to months to a period close to the pre-glitch value. Smaller glitches may be caused by a sudden cracking of the crust of the neutron star and its relaxation to a more spherical shape; larger glitches are probably due to variations in the coupling between the superfluid interior of the star and its independently rotating crust.

1.5 Supermassive Black Holes in Galactic Centers

Black holes are objects so dense that not even light can escape their gravity, and since nothing can travel faster than light, nothing can escape from inside a black hole. On the other hand, a black hole exerts the same force on something far away from it as any other object of the same mass would do. For example, if our Sun were magically crushed until it were about 6 km in size, it would become a black hole, but the Earth would remain in its same orbit.

Where Newton had used the motion of the Moon around the Earth as a guiding example in his work, Einstein used this deviation of Mercury. The central ideas of

general relativity were already in place; at issue was the exact form of the curvature term in the Einstein equation. A precise relativistic model of the Sun's gravitational field was not needed – Einstein used a simple polynomial approximation. Late in the year 1915, he succeeded, and the 43 second lag was explained.

A few weeks later, Einstein, working in Berlin, received a paper from Karl Schwarzschild, an astronomer who, though no longer young, was serving in the German army in Russia. Hospitalized by an illness that soon proved fatal, Schwarzschild had time to discover the desired precise relativistic model, and Schwarzschild spacetime replaced the Newtonian model as the best description of the gravitational field of an isolated spherically symmetrical star. But only a few theorists were familiar with relativity, and significant experimental tests were not possible in Earth-borne laboratories at that time.

In 1963, the British-educated New Zealand physicist Roy Kerr, working at the University of Texas, adopted a shrewd strategy: Bearing in mind that Schwarzschild spacetime has Petrov type D, he did not aim directly at the elusive rotating model, but instead examined an algebraically simple class of type D metric tensors. The long-sought metric appeared. Kerr's minimal one-and-a-half page announcement of his discovery [220] was followed two years later by elaborate detailed calculations [221].

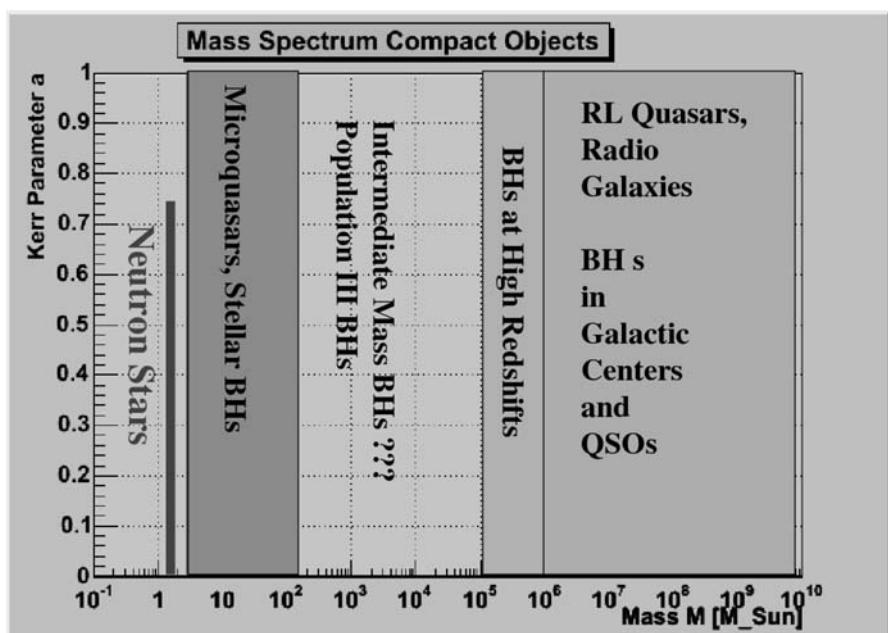


Fig. 1.9. The masses of neutron stars and astrophysical black holes. A black hole has only *two hairs*, the mass and the angular momentum represented by the dimensionless spin parameter a . Neutron stars are clustered around 1.4 solar masses. Stellar black holes are expected to be formed with masses in the range of a few to 100 solar masses. Supermassive black holes in centers of galaxies grow by accretion from black holes formed at high redshifts

Black holes with masses of a million to a few billion times the mass of the Sun are now believed to be the engines that power nuclear activity in galaxies (Fig. 1.10). Active nuclei range from faint, compact radio sources like that in M31 to quasars like 3C 273 that are brighter than the whole galaxy in which they live. Some nuclei fire jets of energetic particles millions of lightyears into space. Almost all astronomers believe that this enormous outpouring of energy comes from the death throes of stars and gas that are falling into the central black hole. This is a very successful explanation of the observations, but until recently, it was seriously incomplete: we had no direct evidence that supermassive black holes exist.

The Hubble Space Telescope provides the best evidence to date of supermassive black holes that lurk in the center of some galaxies. The Space Telescope Imaging Spectrograph (STIS) revealed large orbiting velocities around the nucleus of these galaxies, suggesting a huge mass inside a very small region.

Since the mass of black holes can only grow with time, at least some fraction of nearby galaxies should host such supermassive black holes – like dead quasars with insufficient fuel to trigger the activity in real quasars. For the past 20 years, astronomers have looked for supermassive black holes by measuring rotation and random velocities of stars and gas near galactic centers. If the velocities are large enough, as in the Sombrero Galaxy, then they imply more mass than we see in stars. The most probable explanation is a black hole. About 50 have been found as of the year 2005 (Fig. 1.9). Their masses are in the range expected for nuclear engines,

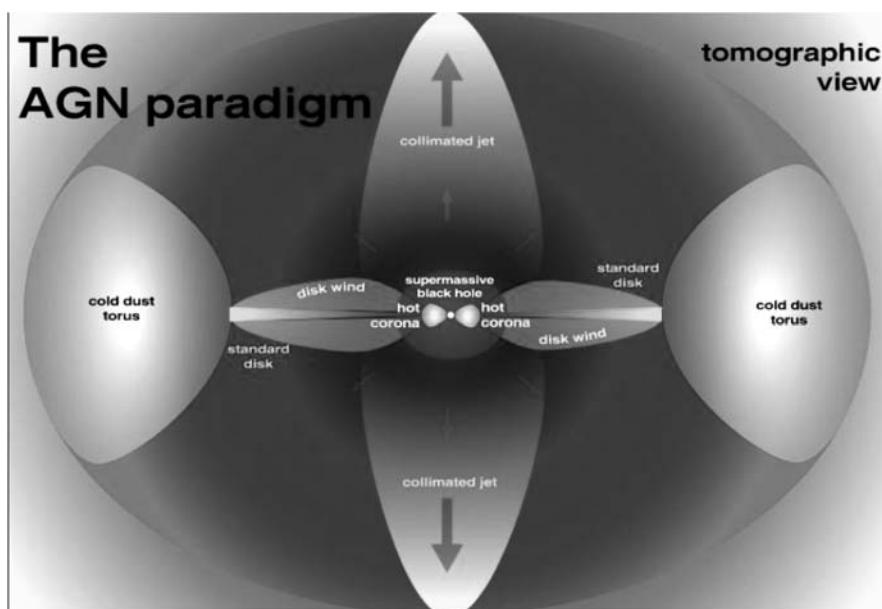


Fig. 1.10. An artist's conception of a supermassive black hole sitting in the center of a galaxy.
Credits: Andreas Müller (ZAH, LSW Heidelberg)

and their numbers are consistent with predictions based on the energy output of quasars.

1.6 Gamma-Ray Bursters

After nearly 30 years of intense debate and scientific inquiry, scientists finally know the answer to the question “Where are the gamma-ray bursts?” and can now move on to answer the question “What causes them?” Recent discoveries in this field by a collection of international astronomers have demonstrated that these bursts are from the most remote parts of the Universe, releasing perhaps as much energy in 10 seconds as the Sun emits in its entire 10-billion-year lifetime.

Gamma-ray bursts (GRBs for short) are brief flashes of high-energy radiation that appear on average about once a day at an unpredictable time from unpredictable directions in the sky (Fig. 1.11). Since their discovery (by accident) in the late 1960s, several thousand bursts have been detected, most of them with BATSE, the Burst and Transient Source Experiment, on board the Compton Gamma Ray Observatory CGRO. Their distribution on the sky is completely uniform (Fig. 1.15). In particular, they do not appear to come from the Milky Way. So where do they come from? This is the question that had kept astronomers busy for several decades, with no apparent resolution in sight.

Gamma-ray bursts (GRBs) were discovered in 1967 by satellites designed to monitor compliance with the atmospheric nuclear test ban treaty. These short-lived outbursts of the most energetic electromagnetic radiation remained one of the biggest mysteries in astrophysics for nearly three decades thereafter. For most of that period, the bursts’ positions in the sky were known only with limited precision, making study of them by ground-based optical and radio telescopes impossible.

With the limited information available, scientists could not determine with confidence even such basic information as the distances from Earth of GRBs. Without such fundamental details, scientists were not able to learn the nature or cause of the bursts. Because of the distance uncertainty, for example, astronomers did not know if GRBs were in our own Solar System, our Galaxy, or in the distant Universe. In this vacuum of knowledge, speculation led to numerous, widely divergent theories that attempted to explain the bursts.

In 1997, the Italian–Dutch satellite BeppoSAX provided ground-based observers with rapid news and more precise positional information about GRBs. On May 8, 1997, BeppoSAX detected a gamma-ray burst, and VLA observers Dale Frail of NRAO and Shri Kulkarni of Caltech discovered radio emission coming from this object on May 13. The mere discovery of radio emission from this gamma-ray burst immediately ruled out some theoretical models at the time. The plot in Fig. 1.12 shows the rapid decay of the X-ray flux coming from this gamma-ray burst event which occurred on May 8, 1997. The X-ray intensity from this source decreased almost one million times in less than one week.

Subsequently, the VLA has detected the radio *afterglow* of several other gamma-ray bursts. While spectroscopy done at optical observatories has shown that GRBs

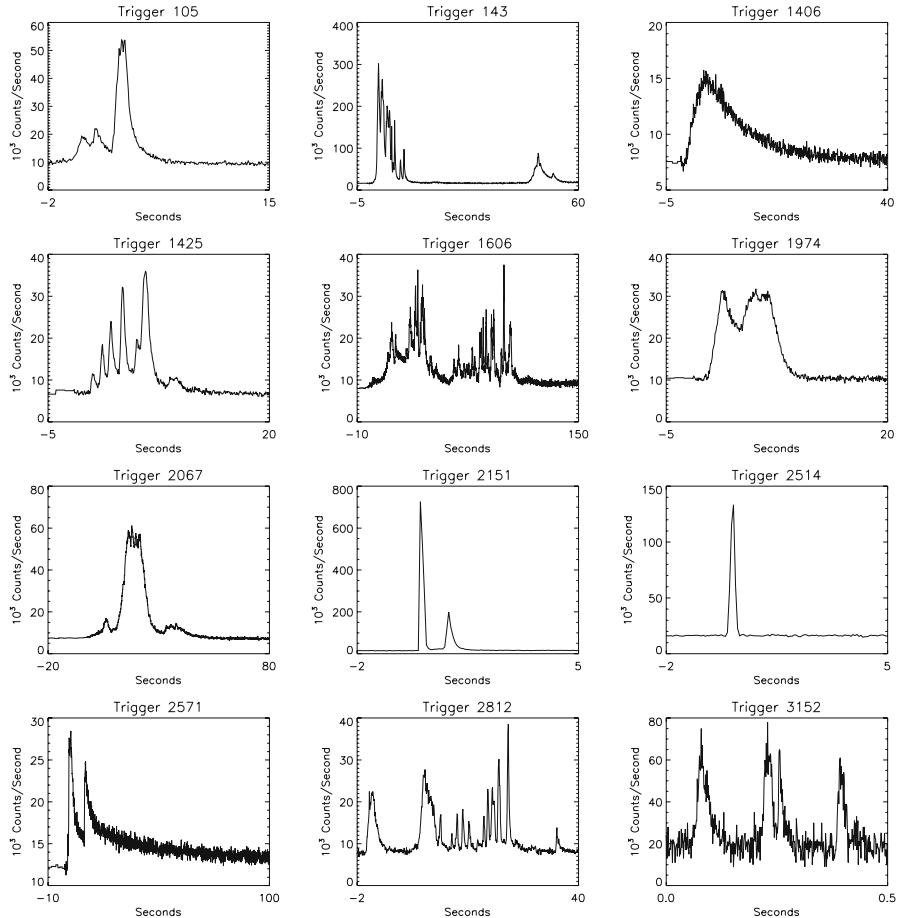


Fig. 1.11. Time profiles for gamma-ray bursts (count rates as a function of time in seconds). Credits: CGRO home page

are at great distances, beyond our own Milky Way Galaxy, VLA observations have revealed the size of the fireball and the speed of its expansion. The May 8, 1997, GRB, for example, was only a tenth of a lightyear across when first detected and expanded at very nearly the speed of light.

The VLA's ability to locate GRBs in the sky with pinpoint precision has helped astronomers at other observatories to locate GRB afterglows that they otherwise might have missed. With the image shown for the GRB of March 29, 1998, the position determined by the VLA was provided to optical and infrared observers, who had failed to find the object, but then, armed with the precise information on its location, found it on images they had already made.

After three decades of mystery, astronomers now know that GRBs, the most violent events in the current Universe, occur in galaxies far from Earth. In addition,

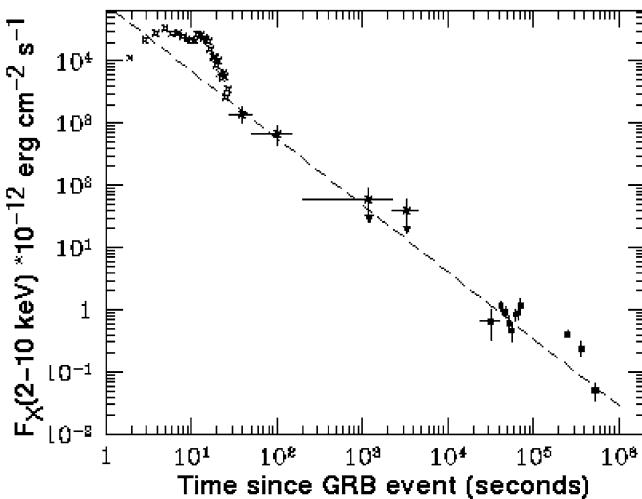


Fig. 1.12. Decay of the GRB970508 flux in X-rays detected by BeppoSAX

the VLA has provided strong evidence that these tremendous explosions occur in dusty areas of those galaxies, where it is likely that young stars are located within the clouds of dust and gas from which they formed. This evidence supports the theory that GRBs result from a **hypernova**, the explosive death of a very massive star that collapses and forms a black hole.

Radio telescopes are the only instruments presently capable of measuring the size of a GRB fireball. In addition, while GRB afterglows fade quickly at other wavelengths, the VLA has been able to follow an afterglow for more than a year, tracking changes in its intensity and other characteristics. These observations indicate the extraordinary importance of radio astronomy for providing information that can be gained in no other way about one of the frontier areas of astrophysics.

Gamma-ray bursts (GRBs) are brief gamma-ray flashes detected with space-based detectors in the range 0.1–100 MeV, with typical photon fluxes of 0.01–100 photons $\text{cm}^{-2} \text{s}^{-1}$ and durations of 0.1–1000 seconds (Fig. 1.13). Their origin is clearly outside the Solar System, and more than a thousand events have been recorded so far. Before there was any firm evidence on the isotropy of classical gamma-ray bursts, the most plausible interpretations involved magnetospheric events on neutron stars (NS) within our Galaxy. However, the remarkable isotropy of these events discovered by the BATSE experiment on the NASA Compton Gamma Ray Observatory (together with the flatter than Newtonian counts) clearly shifts the odds substantially in favor of a cosmological interpretation.

If gamma-ray bursts came from objects in our Galaxy, one would expect to see more of them from the Galactic equator, where most other Galactic objects are found. However, BATSE found that equal numbers of gamma-ray bursts come from all directions (Fig. 1.15).

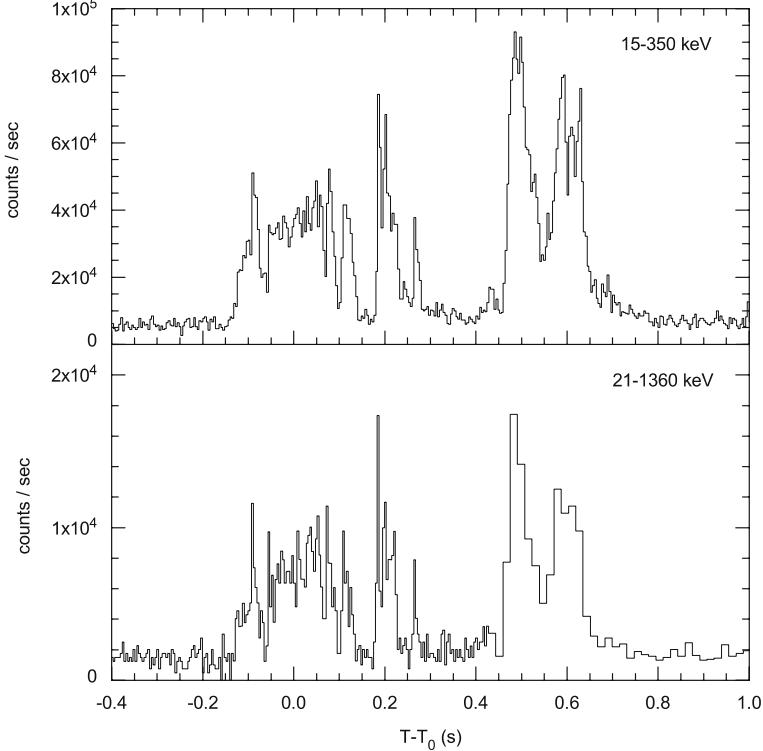


Fig. 1.13. BAT (Swift) and KONUS-Wind light-curves of the short burst GRB060313. The BAT 15–350 keV light-curve (*upper panel*) has at least 20 statistically significant peaks with FWHMs in the 5–15 ms range. The KONUS-Wind 21–1360 keV light-curve (*lower panel*) also exhibits several multipeaked pulses with a total duration of approximately 0.8 s. Figure adapted from Roming et al. [347]

In principle, the isotropy could be interpreted in terms of either (i) a cosmological distribution similar to that of the distant galaxies and clusters, i.e. hundreds and thousands of Mpc, (ii) a distribution in an “extended halo” of our galaxy, which is so large that the small dipole moment associated with our off-center location is not noticeable (i.e. greater than $\simeq 200$ kpc), or (iii) a “galactic disk” distribution, where objects are sufficiently faint that they are detectable only out to distances smaller than the width of the disk (few kpc). The “galactic disk” model has difficulty in explaining the large number of events (a few per day) occurring within a few kpc, and the dipole and quadrupole moment of the spatial distribution appear to rule out such an origin. The “extended halo” option may satisfy (just) the dipole and quadrupole observational restrictions, but the physical origin of the bursts and the number of sources at such large distances is not straightforward to explain. On the other hand, the “cosmological” interpretation does have at least two rather plausible energy sources: either NS-NS (or NS-black-hole) binary mergers (e.g.

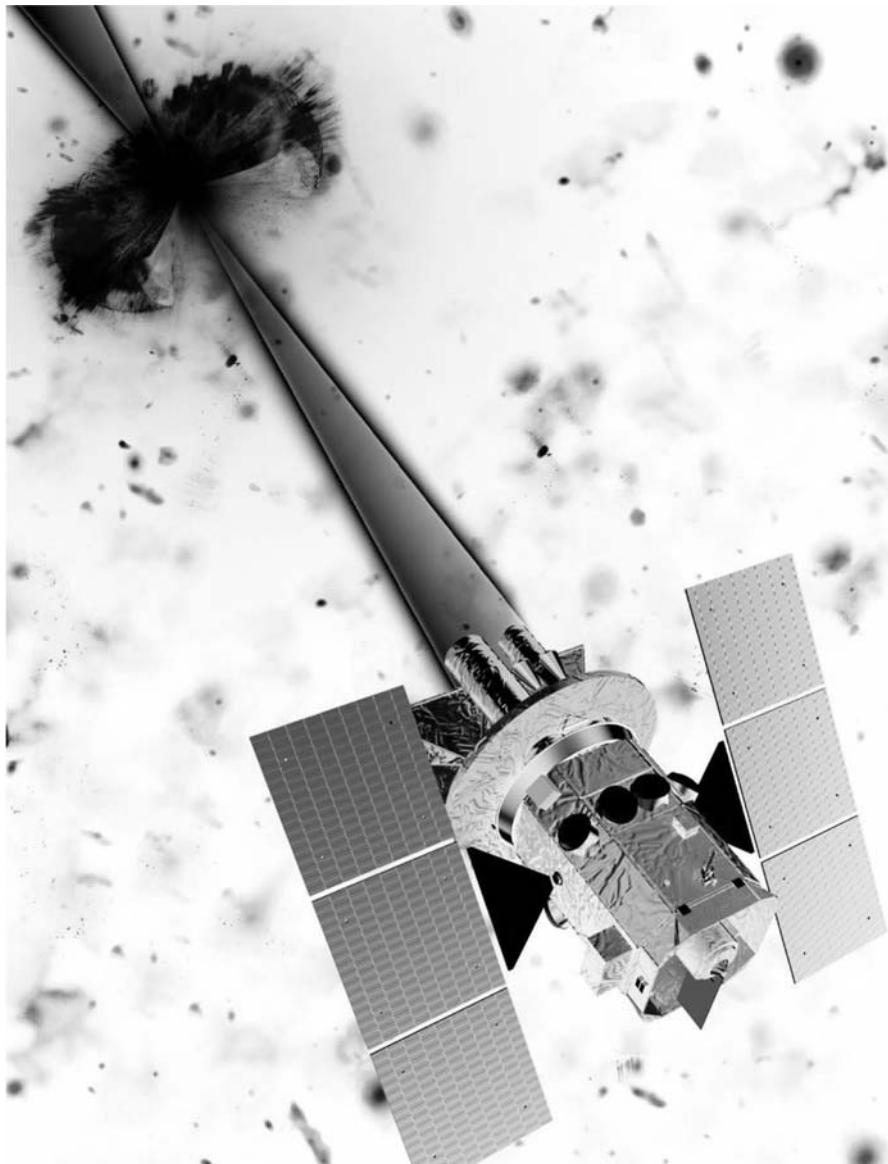


Fig. 1.14. The Swift satellite in observing a gamma-ray burst emitted from the jets launched by a hypernova. Credit: Spectrum-Astro

binary pulsars merging under the effect of gravitational wave energy losses), or else “failed supernova” events (where a star undergoes core collapse to a NS but with much reduced optical display). Either of these should occur with a frequency of 10^{-5} per galaxy per year, and produces $10^{50}\text{--}10^{51}$ ergs, detectable out to redshifts

2704 BATSE Gamma-Ray Bursts

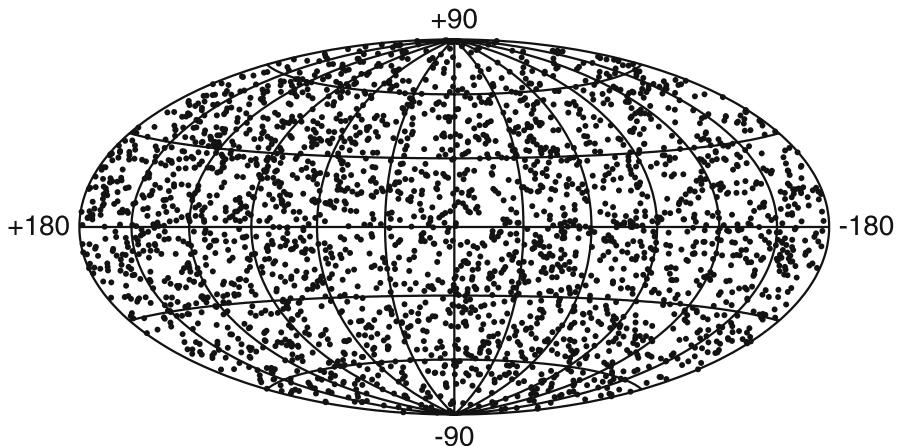


Fig. 1.15. Distribution of GRBs on the sky as observed by BATSE on CGRO. Source: CGRO home page

of order unity, so that the typical frequency and fluence is easily explained. More importantly, the discovery in February 1997 of GRB afterglows and counterparts gives strong support to the cosmological origin.

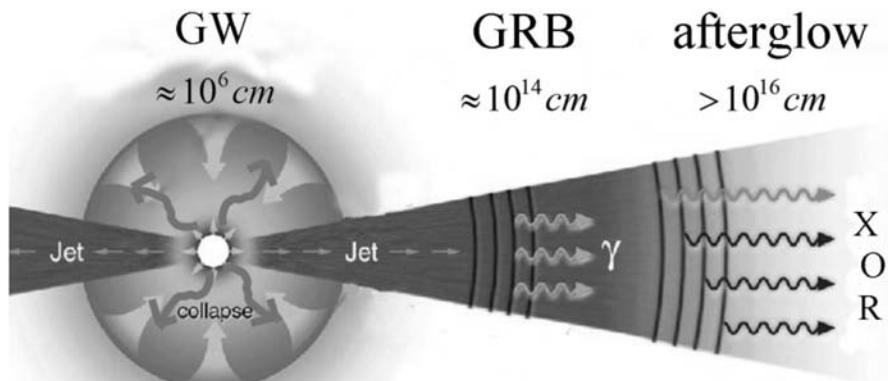


Fig. 1.16. Schematics of GRB models. The collapse towards a black hole produces large amounts of electron–positron pairs which escape as relativistic jets in polar directions. Initially, the plasma is optically thick, and only at distances of the order of astronomical units, can gamma-rays escape. These jets drive a shock-front which is visible in the afterglows

One model for GRBs is a binary pair of neutron stars, with their orbital separation ground down by billions of years of gravitational radiation, finally merging to form a black hole. In the process of merging the neutron stars tear each other apart, forming an accretion disk and jets (Fig. 1.16). A straightforward prediction of cosmological models is that, if GRB are standard candles, one would expect the weaker fluence bursts (which presumably are farther) to have longer durations due to cosmological time dilation. Such an effect has been recently reported. However, the duration of the burst can depend on intrinsic properties of the source, so that this cosmological signature may be smeared out by details of the source physics. For cosmological GRBs, a very interesting prediction is that they should be accompanied by gravitational wave bursts of energy comparable to a solar rest mass. These would be detectable at the rate of several per year with coincidence measurements from two advanced versions of the proposed LIGO or VIRGO detectors. Such measurements might also distinguish between failed supernova events or compact binary mergers, through their wave profile. One may also obtain valuable information concerning early star formation, through limits on the typical redshift derived from the counts of events as a function of the fluence, and it may be possible to derive limits on the GRB luminosity distribution.

Problems

- 1.1.** Derive an estimate for the number of white dwarf stars in a globular cluster, based on the initial mass function (IMF).
- 1.2.** Derive an estimate for the number of neutron stars in our Galaxy, based on the observed supernova rate.
- 1.3.** Radio astronomers determine the period P and its derivative \dot{P} by means of pulsar timing. Determine the energy-loss rate for a rotating neutron star with period P in seconds and \dot{P} in units of $10^{-15} \text{ s s}^{-1}$.
- 1.4.** What is the fluence for gamma-ray bursts? Find a fluence histogram for GRBs on the Web.
- 1.5.** Discuss the methods to derive the masses of black holes in galactic centers.
- 1.6.** Explore the meaning of the “Magorrian relation” between the black hole mass and the stellar velocity dispersion for galaxies. Find out correlations between the black hole mass and the stellar mass of the bulge in galaxies.

2 Gravity of Compact Objects

The gravity of compact objects requires a description of gravitational fields much beyond the Newtonian picture. In this chapter, we give a short overview for the most important concepts and methods of general relativity. This does not replace a thorough study of Einstein's theory. This marvellous theory is explained in many classical textbooks; see for example the books by Misner, Thorne and Wheeler [10], Schutz [14], Carroll [2], or Straumann [18]. Since recent research on compact objects goes much beyond a simple stationary description of gravitational fields, we also give a short introduction to the concepts of the 3+1 split of Einstein's equations, which is now the basis of numerical treatments of Einstein's field equations. Simulations for the merging of two black holes or two neutron stars are based on these techniques. The 3+1 technique is now a very powerful method, which can also be implemented in deriving, e.g. the field equations for rapidly rotating compact objects, such as neutron stars and black holes.

2.1 Geometric Concepts and General Relativity

In 1915 Albert Einstein published a geometrical theory of gravitation [141]: the general theory of relativity. He presented a fundamentally new description of gravity in the sense that the relative acceleration of particles is not viewed as a consequence of gravitational forces, but results from the curvature of the spacetime in which the particles are moving. As long as no nongravitational forces act on a particle, it is always moving on a “straight line.” If we consider curved manifolds, there is still a concept of straight lines which are called geodesics, but these will not necessarily have the properties we intuitively associate with straight lines from our experience in flat Euclidean geometry. It is, for example, a well known fact that two distinct straight lines in two-dimensional flat geometry will intersect each other exactly once, unless they are parallel, in which case they do not intersect each other at all. These ideas result from the fifth Euclidean postulate of geometry, which plays a special role in the formulation of geometry.

It is a well known fact that one needs to impose it separately from the first four Euclidean postulates in order to obtain flat Euclidean geometry. It was not realized until the work of Gauss, Lobachevsky, Bolyai and Riemann in the 19th century that the omission of the fifth postulate leads to an entirely new class of non-Euclidean geometries in curved manifolds. A fundamental feature of non-Euclidean

geometry is that straight lines in curved manifolds can intersect each other more than once and correspondingly diverge from and converge towards each other several times.

In order to illustrate how these properties give rise to effects we commonly associate with forces such as gravitation, we consider two observers on the Earth's surface, say one in Heidelberg, Germany, and one in Vienna, Austria. We assume that these two observers start moving due south in "straight lines" as for example guided by an idealized compass exactly pointing towards the south pole. If we follow their separate paths we will discover exactly the ideas outlined above. As long as both observers are in the northern hemisphere the proper distance between them will increase and reach a maximum when they reach the equator. From then on they will gradually approach each other and their paths will inevitably cross at the south pole.

In the framework of Newtonian physics, the observers will attribute the relative acceleration of their positions to the action of a force. It is clear, however, that no force is acting in the east–west direction on either observer at any stage of their journey. In a geometric description, the relative movement of the observers finds a qualitatively new interpretation in terms of the curvature of the manifold they are moving in, the curvature of the Earth's surface. With the development of general relativity, Einstein provided the exact mathematical foundation for applying these ideas to the forces of gravitation in four-dimensional space-time.

One may ask why such a geometrical interpretation has only been developed for gravitation. Or in other words: which feature distinguishes gravitation from the other three fundamental interactions? The answer lies in the gravitational charge, the mass. It is a common observation that the gravitational mass m_G which determines the coupling of a particle to the gravitational field is virtually identical to the inertial mass m_I which describes the particle's kinematic reaction to an external force. High precision experiments have been undertaken to measure the difference between these two types of masses. All these results are compatible with the assumption that the masses are indeed equal. The mass will therefore drop out of the Newtonian equations governing the dynamics of a particle subject exclusively to gravitational forces $m a = GmM/r^2$, where a is the acceleration of the particle, G the gravitational constant, M the mass of an external source and r the distance from this source. The particle mass m can be factored out so that the movement of the particle is described in purely kinematic terms. The redundancy of the concept of a gravitational force is naturally incorporated into a geometric theory of gravity such as general relativity.

It is important to note that this behavior distinguishes gravity from the other fundamental interactions which are associated with different types of charges, such as electric charge in the case of electromagnetic interaction. It is not obvious how and whether it is possible to obtain similar geometric formulations for the electromagnetic, weak and strong interaction. The unification of these three fundamental forces with gravity in the framework of quantum theory is one of the important areas of ongoing research.

2.2 The Basic Principles of General Relativity

2.2.1 Einstein's Equivalence Principle and Metricity

The principle of equivalence has historically played an important role in the development of gravitation theory. Newton regarded this principle as such a cornerstone of mechanics that he devoted the opening paragraph of the *Principia* to it. In 1907, Einstein used the principle as a basic element of general relativity. We now regard the principle of equivalence as the foundation, not of Newtonian gravity or of GR, but of the broader idea that spacetime is curved. One elementary equivalence principle is the kind Newton had in mind when he stated that the property of a body called “mass” is proportional to the “weight,” and is known as the **weak equivalence principle** (WEP). An alternative statement of WEP is that the trajectory of a freely falling body (one not acted upon by such forces as electromagnetism and too small to be affected by tidal gravitational forces) is independent of its internal structure and composition. In the simplest case of dropping two different bodies in a gravitational field, WEP states that the bodies fall with the same acceleration (this is often termed the Universality of Free Fall).

A more powerful and far-reaching equivalence principle is known as the **Einstein equivalence principle** (EEP). It states that [419]:

1. **WEP is valid.**
2. **The outcome of any local nongravitational experiment is independent of the velocity of the freely falling reference frame in which it is performed.**
3. **The outcome of any local nongravitational experiment is independent of where and when in the Universe it is performed.**

The second piece of EEP is called local Lorentz invariance (LLI), and the third piece is called local position invariance (LPI).

For example, a measurement of the electric force between two charged bodies is a local nongravitational experiment; a measurement of the gravitational force between two bodies (Cavendish experiment) is not.

The Einstein equivalence principle is the heart and soul of gravitational theory, for it is possible to argue convincingly that if EEP is valid, then gravitation must be a **curved spacetime** phenomenon, in other words, the effects of gravity must be equivalent to the effects of living in a curved spacetime. As a consequence of this argument, the only theories of gravity that can embody EEP are those that satisfy the postulates of **metric theories of gravity**, which are:

1. **Spacetime is endowed with a symmetric metric.**
2. **The trajectories of freely falling bodies are geodesics of that metric.**
3. **In local freely falling reference frames, the nongravitational laws of physics are those written in the language of special relativity.**

The argument that leads to this conclusion simply notes that, if EEP is valid, then in local freely falling frames, the laws governing experiments must be independent of the velocity of the frame (local Lorentz invariance), with constant values for the

various atomic constants (in order to be independent of location). The only laws we know of that fulfill this are those that are compatible with special relativity, such as Maxwell's equations of electromagnetism. Furthermore, in local freely falling frames, test bodies appear to be unaccelerated, in other words they move on straight lines; but such “locally straight” lines simply correspond to “geodesics” in a curved spacetime.

General relativity is a metric theory of gravity, but then so are many others, including the Brans–Dicke theory. Neither, in this narrow sense, is superstring theory, which, while based fundamentally on a spacetime metric, introduces additional fields (dilatons, moduli) that can couple to material stress–energy in a way that can lead to violations, say, of WEP. Therefore, the notion of curved spacetime is a very general and fundamental one, and therefore it is important to test the various aspects of the Einstein Equivalence Principle thoroughly.

A direct test of WEP is the comparison of the acceleration of two laboratory-sized bodies of different composition in an external gravitational field. If the principle were violated, then the accelerations of different bodies would differ. The simplest way to quantify such possible violations of WEP in a form suitable for comparison with experiment is to suppose that for a body with inertial mass m_I , the passive gravitational mass m_P is no longer equal to m_I , so that in a gravitational field g , the acceleration is given by

$$m_I a = m_P g . \quad (2.1)$$

Now the inertial mass of a typical laboratory body is made up of several types of mass–energy: rest energy, electromagnetic energy, weak-interaction energy, and so on. If one of these forms of energy contributes to m_P differently than it does to m_I , a violation of WEP would result. One could then write

$$m_P = m_I + \sum_A \eta_A E^A / c^2 , \quad (2.2)$$

where E^A is the internal energy of the body generated by interaction A, and η_A is a dimensionless parameter that measures the strength of the violation of WEP induced by that interaction, and c is the speed of light. A measurement or limit on the fractional difference in acceleration a_1 and a_2 measured between two bodies then yields a quantity called the **Eötvös ratio** defined as

$$\eta = \frac{2|a_1 - a_2|}{|a_1 + a_2|} = \sum_A \eta_A \left(\frac{E_1^A}{m_{I,1}c^2} - \frac{E_2^A}{m_{I,2}c^2} \right) . \quad (2.3)$$

Many high-precision Eötvös-type experiments have been performed, from the pendulum experiments of Newton, Bessel and Potter, to the classic torsion-balance measurements of Eötvös, Dicke, Braginsky and their collaborators. In the modern torsion-balance experiments, two objects of different composition are connected by a rod or placed on a tray and suspended in a horizontal orientation by a fine wire. If the gravitational acceleration of the bodies differs, there will be a torque induced

on the suspension wire, related to the angle between the wire and the direction of the gravitational acceleration g . If the entire apparatus is rotated about some direction with angular velocity ω , the torque will be modulated with period $2\pi/\omega$. In the experiments of Eötvös and his collaborators, the wire and g were not quite parallel because of the centripetal acceleration on the apparatus due to the Earth's rotation; the apparatus was rotated about the direction of the wire. In the Dicke and Braginsky experiments, g was that of the Sun, and the rotation of the Earth provided the modulation of the torque at a period of 24 hr. Beginning in the late 1980s, numerous experiments were carried out primarily to search for a “fifth force,” but their null results also constituted tests of WEP. In the “free-fall Galileo experiment” performed at the University of Colorado, the relative free-fall acceleration of two bodies made of uranium and copper was measured using a laser interferometric technique. The “Eöt-Wash” experiments carried out at the University of Washington used a sophisticated torsion balance tray to compare the accelerations of various materials toward local topography on Earth, movable laboratory masses, the Sun and the galaxy, and have recently reached levels of 4×10^{-13} . The resulting upper limits on η are summarized in Fig. 2.1.

Gravitational Redshift

The gravitational redshift is one of the most prominent consequences of EEP. Consider two labs, a distance h apart moving with constant acceleration a . At time t_0 , the first experiment emits a photon of wavelength λ_0 . The two experiments remain a constant distance apart, so that the photon reaches the leading apparatus after a time $\Delta t = h/c$ in the reference frame of the experiments. In this time, the apparatus have picked up an additional velocity $\Delta v = a \Delta t = ah/c$. Therefore, the photon reaching the leading apparatus will be redshifted by the conventional Doppler effect

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta v}{c} = \frac{ah}{c^2}. \quad (2.4)$$

According to the EEP, the same situation should happen in a uniform gravitational field. So we can imagine a tower of height z on the Earth surface with a_g the strength of the gravitational field. This situation is supposed to be indistinguishable from the previous one in a lab on top of the tower, where the photon from the ground will be detected. Therefore, a photon emitted from the ground with wavelength λ_0 should be redshifted by the amount

$$\frac{\Delta\lambda}{\lambda} = \frac{a_g h}{c^2}. \quad (2.5)$$

This is the famous **gravitational redshift**. This effect is a direct consequence of the EEP, not of the details of general relativity! It has been verified for the first time by Pound and Rebka in 1960. They used the Mössbauer effect to measure the change in frequency of gamma-rays as they travelled from the ground to the top of Jefferson labs in Harvard.

TESTS OF THE WEAK EQUIVALENCE PRINCIPLE

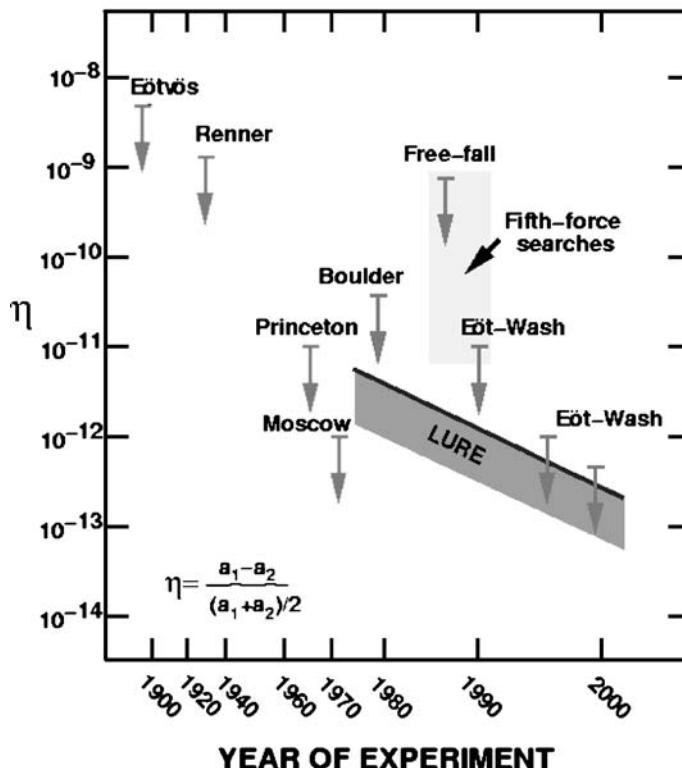


Fig. 2.1. Selected tests of the weak equivalence principle, showing bounds on η , which measures fractional difference in acceleration of different materials or bodies. The free-fall and Eöt-Wash experiments were originally performed to search for a fifth force. The *shaded band* shows current bounds on η for gravitating bodies from lunar laser ranging (LURE). Credits: C. Will [419]

This formula for the redshift can be stated in terms of the gravitational potential U , where $a_g = \nabla U$ ¹. From this we obtain the redshift

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{c^2} \int \partial_z U dz = \frac{\Delta U}{c^2}. \quad (2.6)$$

The principle of local position invariance, the third part of EEP, can be tested by the gravitational redshift experiment, the first experimental test of gravitation

¹ The sign is changed with respect to the usual convention, since we are thinking of a_g as the acceleration of the reference frame, and not of a particle with respect to this reference frame.

proposed by Einstein. Despite the fact that Einstein regarded this as a crucial test of GR, we now realize that it does not distinguish between GR and any other metric theory of gravity, but is only a test of EEP. A typical gravitational redshift experiment measures the frequency or wavelength shift $z = \Delta\nu/\nu = -\Delta\lambda/\lambda$ between two identical frequency standards (clocks) placed at rest at different heights in a static gravitational field. If the frequency of a given type of atomic clock is the same when measured in a local, momentarily comoving freely falling frame (Lorentz frame), independent of the location or velocity of that frame, then the comparison of frequencies of two clocks at rest at different locations boils down to a comparison of the velocities of two local Lorentz frames, one at rest with respect to one clock at the moment of emission of its signal, the other at rest with respect to the other clock at the moment of reception of the signal. The frequency shift is then a consequence of the first-order Doppler shift between the frames. The structure of the clock plays no role whatsoever. The result is a shift

$$z = \Delta U/c^2, \quad (2.7)$$

where ΔU is the difference in the Newtonian gravitational potential between the receiver and the emitter. If LPI is not valid, then it turns out that the shift can be written

$$z = (1 + \alpha) \Delta U/c^2, \quad (2.8)$$

where the parameter α may depend upon the nature of the clock whose shift is being measured.

2.2.2 Metric Theories of Gravity

In order to formalize the ideas mentioned in the previous section, general relativity views spacetime as a four-dimensional manifold equipped with a metric $g_{\alpha\beta}$ of Lorentzian signature where the Greek indices range from 0 to 3. At any given point in the manifold the signature enables one to distinguish between time-like, space-like and null directions. The metric further induces a whole range of higher level geometric concepts on the manifold. It defines a scalar product between vectors which leads to the measurement of length and the idea of orthogonality. From the metric and its derivatives one can derive a connection on the manifold which facilitates the definition of a covariant derivative. The notion of a derivative is more complicated in a curved manifold than in the common case of flat geometry and Cartesian coordinates because the basis vectors will in general vary from point to point in the manifold. It is therefore no longer possible to identify the derivative of a tensor with the derivative of its components. Instead one obtains extra terms involving the derivatives of the basis vectors. In terms of a covariant derivative, these terms are represented by the connection. In general relativity one uses a metric-compatible connection defined by

$$\Gamma_{\alpha\beta}^\mu = \frac{1}{2} g^{\mu\varrho} (\partial_\alpha g_{\beta\varrho} + \partial_\beta g_{\alpha\varrho} - \partial_\varrho g_{\alpha\beta}), \quad (2.9)$$

where the Einstein summation convention, according to which one sums over repeated upper and lower indices, has been used. These connection coefficients are also known as the Christoffel symbols and define a covariant derivative of tensors of arbitrary rank by

$$\nabla_\varrho T^\alpha_\beta = \partial_\varrho T^\alpha_\beta + \Gamma^\alpha_{\varrho\mu} T^\mu_\beta - T^\alpha_\mu \Gamma^\mu_{\varrho\beta}. \quad (2.10)$$

So for each upper index one adds a term containing the connection coefficients and for each lower index a corresponding term is subtracted. With the definition of a covariant derivative we can finally write down the exact definition of a “straight line” in a curved manifold. A geodesic is defined as the integral curve of a vector field U which is parallel transported along itself

$$(U^\alpha \nabla_\alpha) U = 0. \quad (2.11)$$

Based on the covariant derivative we can also give a precise definition of curvature. If we use a coordinate basis, this definition can be shown to imply that for any vector field U

$$\nabla_\varrho \nabla_\sigma U^\alpha - \nabla_\sigma \nabla_\varrho U^\alpha = R^\alpha_{\beta\varrho\sigma} U^\beta, \quad (2.12)$$

which is commonly interpreted by saying that a vector U is changed by being parallel transported around a closed loop unless the curvature vanishes.

In order to describe the effect of the matter distribution on the geometry of spacetime, one defines the Ricci tensor as the contraction of the Riemann tensor $R_{\beta\delta} = R^\alpha_{\beta\alpha\delta}$, where again the Einstein summation convention for repeated indices has been used. Geometry and matter are then related by

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = \kappa T_{\alpha\beta}, \quad (2.13)$$

where R is the Ricci scalar, $R = R^\alpha_\alpha$, and $T_{\alpha\beta}$ the energy momentum tensor. The interaction between the matter distribution and the geometry of spacetime can be summed up in the words of Misner, Thorne and Wheeler: *Space acts on matter, telling it how to move. In turn, matter reacts back on space, telling it how to curve.*

The Strong Equivalence Principle (SEP)

In any metric theory of gravity, matter and nongravitational fields respond only to the spacetime metric g . In principle, however, there could exist other gravitational fields besides the metric, such as scalar fields, vector fields, and so on. If, by our strict definition of metric theory, matter does not couple to these fields, what can their role in gravitation theory be? Their role must be that of mediating the manner in which matter and nongravitational fields generate gravitational fields and produce the metric; once determined, however, the metric alone acts back on the matter in the manner prescribed by EEP.

What distinguishes one metric theory from another, therefore, is the number and kind of gravitational fields it contains in addition to the metric, and the equations that determine the structure and evolution of these fields. From this viewpoint, one can divide all metric theories of gravity into two fundamental classes: “purely dynamical” and “prior-geometric” [419].

By “purely dynamical metric theory” we mean any metric theory whose gravitational fields have their structure and evolution determined by coupled partial differential field equations. In other words, the behavior of each field is influenced to some extent by a coupling to at least one of the other fields in the theory. By “prior geometric” theory, we mean any metric theory that contains “absolute elements,” fields or equations whose structure and evolution are given *a priori*, and are independent of the structure and evolution of the other fields of the theory. These “absolute elements” typically include flat background metrics η , cosmic time coordinates t , and algebraic relationships among otherwise dynamical fields.

General relativity is a purely dynamical theory, since it contains only one gravitational field, the metric itself, and its structure and evolution are governed by partial differential equations (Einstein’s equations). Brans–Dicke theory and its generalizations are purely dynamical theories, too; the field equation for the metric involves the scalar field (as well as the matter as source), and that for the scalar field involves the metric.

By discussing metric theories of gravity from this broad point of view, it is possible to draw some general conclusions about the nature of gravity in different metric theories, conclusions that are reminiscent of the Einstein equivalence principle, but that are subsumed under the name “strong equivalence principle.”

Consider a local, freely falling frame in any metric theory of gravity. Let this frame be small enough that inhomogeneities in the external gravitational fields can be neglected throughout its volume. On the other hand, let the frame be large enough to encompass a system of gravitating matter and its associated gravitational fields. The system could be a star, a black hole, the Solar System or a Cavendish experiment. Call this frame a “quasilocal Lorentz frame.” To determine the behavior of the system, we must calculate the metric. The computation proceeds in two stages. First we determine the external behavior of the metric and gravitational fields, thereby establishing boundary values for the fields generated by the local system, at a boundary of the quasilocal frame “far” from the local system. Second, we solve for the fields generated by the local system. But because the metric is coupled directly or indirectly to the other fields of the theory, its structure and evolution will be influenced by those fields, and in particular by the boundary values taken on by those fields far from the local system. This will be true, even if we work in a coordinate system in which the asymptotic form of g in the boundary region between the local system and the external world is that of the Minkowski metric. Thus the gravitational environment, in which the local gravitating system resides, can influence the metric generated by the local system via the boundary values of the auxiliary fields. Consequently, the results of local gravitational experiments may depend on the location and velocity of the frame relative to the external environment.

Of course, local nongravitational experiments are unaffected, since the gravitational fields they generate are assumed to be negligible, and since those experiments couple only to the metric, whose form can always be made locally Minkowskian at a given spacetime event. Local gravitational experiments might include Cavendish experiments, measurement of the acceleration of massive self-gravitating bodies, studies of the structure of stars and planets, or analyses of the periods of “gravitational clocks.” We can now make several statements about different kinds of metric theories.

- A theory which contains only the metric g yields local gravitational physics which is independent of the location and velocity of the local system. This follows from the fact that the only field coupling the local system to the environment is g , and it is always possible to find a coordinate system in which g takes the Minkowski form at the boundary between the local system and the external environment. Thus the asymptotic values of g are constants independent of location, and are asymptotically Lorentz invariant, thus independent of velocity. General relativity is an example of such a theory.
- A theory, which contains the metric g and dynamical scalar fields, yields local gravitational physics, which may depend on the location of the frame but which is independent of the velocity of the frame. This follows from the asymptotic Lorentz invariance of the Minkowski metric and of the scalar fields, but now the asymptotic values of the scalar fields may depend on the location of the frame. An example is Brans–Dicke theory, where the asymptotic scalar field determines the effective value of the gravitational constant, which can thus vary as the scalar field varies. On the other hand, a form of velocity dependence in local physics can enter indirectly if the asymptotic values of the scalar field vary with time cosmologically. Then the rate of variation of the gravitational constant could depend on the velocity of the frame.
- A theory which contains the metric g and additional dynamical vector or tensor fields or prior-geometric fields yields local gravitational physics which may have both location and velocity-dependent effects.

These ideas can be summarized in the **strong equivalence principle (SEP)**, which states that

1. WEP is valid for self-gravitating bodies as well as for test bodies.
2. The outcome of any local test experiment is independent of the velocity of the (freely falling) apparatus.
3. The outcome of any local test experiment is independent of where and when in the Universe it is performed.

The distinction between SEP and EEP is the inclusion of bodies with self-gravitational interactions (planets, stars) and of experiments involving gravitational forces (Cavendish experiments, gravimeter measurements). Note that SEP contains EEP as the special case in which local gravitational forces are ignored.

2.3 Basic Calculus on Manifolds

The above discussion suggests to replace the flat Minkowskian spacetime by means of a curved manifold. In this section, I give a short outline about structures of **manifolds**. A manifold is one of the most fundamental concepts in mathematical physics. The notion of a manifold captures the idea of a curved space, which is however locally just flat. In fact, the entire manifold is glued together by local patches.

A manifold is a topological space which is locally Euclidean (i.e. around every point, there is a neighborhood which is topologically the same as the open unit ball in \mathbf{R}^n). To illustrate this idea, consider the ancient belief that the Earth was flat as contrasted with the modern evidence that it is round. This discrepancy arises essentially from the fact that on the small scales that we see, the Earth does indeed look flat (although the Greeks did notice that the last part of a ship to disappear over the horizon was the mast). In general, any object which is nearly “flat” on small scales is a manifold, and so manifolds constitute a generalization of objects we could live on in which we would encounter the round/flat Earth problem, as first codified by Poincaré. More formally, **any object that can be charted is a manifold**. For the mathematical details, see any textbook on manifolds.

2.3.1 Tensors and Forms on Manifolds

Tangent Vectors

With the definition of a function $f(x^1(t), \dots, x^n(t))$ on a curve γ , parametrized by t , in a manifold M we now consider its derivative

$$\begin{aligned} \left(\frac{\partial f}{\partial t} \right)_{\gamma(t)|_{t=t_0}} &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left[f(\gamma(t_0 + \epsilon)) - f(\gamma(t_0)) \right] \\ &= \sum_{j=1}^n \frac{dx^j(t)}{dt} \Big|_{t_0} \left(\frac{\partial f}{\partial x^j} \right)_{\gamma(t_0)} = \left(\frac{dx^j}{dt} \frac{\partial f}{\partial x^j} \right)_{t_0}. \end{aligned} \quad (2.14)$$

In the last step we use the summation convention. Since the function f was arbitrary, we have

$$\frac{\partial}{\partial t} = \frac{dx^j}{dt} \frac{\partial}{\partial x^j}. \quad (2.15)$$

Thus the partials do in fact form a good basis for the vector space of directional derivatives, which we identify with the tangent space.

By considering various curves γ passing through a given point p we can define a linear vector space at p consisting of linear combinations of the coordinate derivatives $\partial/\partial x^j$ in the form

$$X = X^j \frac{\partial}{\partial x^j}, \quad (2.16)$$

where the X^j 's are any set of n numbers. These **tangent vectors** arise by considering the curves γ defined by

$$x^j(t) = x^j(p) + X^j t \quad (2.17)$$

for t in some interval $-\epsilon < t < \epsilon$.

Finally, the definition of X by

$$X.f = X^j \frac{\partial f}{\partial x^j} \equiv X^j f_{,j} \quad (2.18)$$

clearly satisfies the Leibniz rule when operating on products of functions

$$X.(f \circ g)|_{\gamma(t)} = (f X.g + g X.f)|_{\gamma(t)}. \quad (2.19)$$

Tangent vectors may indeed be considered as directional derivatives, \circ is the composition operator.

The tangent vectors at p form a linear vector space of the reals spanned by the coordinate derivatives, since

$$(\alpha X + \beta Y).f = \alpha(X.f) + \beta(Y.f) \quad (2.20)$$

is satisfied for all vectors X and Y , numbers α and β and functions f . In addition, the vectors $(\partial/\partial x^j)_p$ are linearly independent.

The space of all tangent vectors (as also called contravariant vectors) to an n -dimensional manifold M at p is denoted by $T_p M$, or simply T_p . This space, which may be visualized as the set of all directions at p , is called the **tangent space** at p .

Vielbeine or Tetrads

Instead of a basis determined by local coordinates, $\partial/\partial x^j$, we may choose any other n linearly independent vectors \mathbf{e}_a ($a = 1, \dots, n$). There must then exist linear relations of the form

$$\mathbf{e}_a = \Phi_a^k \frac{\partial}{\partial x^k}, \quad (2.21)$$

where the determinant of the matrix formed by Φ_a^k must be nonzero. The inverse relation is then given by

$$\frac{\partial}{\partial x^k} = \Phi_k^b \mathbf{e}_b, \quad (2.22)$$

where Φ_k^b is the inverse of the matrix Φ_a^k , i.e. it satisfies

$$\Phi_a^k \Phi_j^a = \delta_j^k, \quad \Phi_a^k \Phi_k^b = \delta_a^b. \quad (2.23)$$

Such a basis is called a vielbein, in analogy to the four-dimensional vierbeins. Given a vielbein basis, we can express any tangent vector at p in the form

$$X = X^a \mathbf{e}_a. \quad (2.24)$$

The X^a 's are now the components of the vector X relative to the vielbein basis.

One-Forms

A one-form ω at p is a linear mapping of the tangent space $T_p M$ onto the reals, $\omega : T_p M \rightarrow \mathbf{R}$. This means, given any tangent vector X at p , the one-form ω associates uniquely with it a number $\omega(X)$ which is also written as

$$\omega(X) = \langle \omega, X \rangle . \quad (2.25)$$

The required linearity of the map is expressed by the relation

$$\langle \omega, \alpha X + \beta Y \rangle = \alpha \langle \omega, X \rangle + \beta \langle \omega, Y \rangle \quad (2.26)$$

for any two tangent vectors X and Y and any reals α and β . We also define multiplications of forms and sums of forms by means of

$$(\alpha\omega)(X) = \alpha \langle \omega, X \rangle, \quad (\omega + \pi)(X) = \langle \omega, X \rangle + \langle \pi, X \rangle . \quad (2.27)$$

By these rules, one-forms also span a vector space which we denote by $T_p^* M$. It is called the **cotangent space at p and the dual of the tangent space**. For this reason, one-forms are also called cotangent vectors (or, covariant vectors).

A basis for $T_p^* M$, associated with the vielbein basis \mathbf{e}_a for $T_p M$, is provided by the one-form basis \mathbf{e}^a ($a = 1, \dots, n$) which map any tangent vector $X = X^a \mathbf{e}_a$ to its components

$$\mathbf{e}^a(X) = \langle \mathbf{e}^a, X^b \mathbf{e}_b \rangle = X^a . \quad (2.28)$$

From this we have the duality

$$\mathbf{e}^a(\mathbf{e}_b) = \langle \mathbf{e}^a, \mathbf{e}_b \rangle = \delta_b^a . \quad (2.29)$$

Similar to tangent vectors, we may also express each one-form ω in terms of a dual vielbein basis

$$\omega = \omega_a \mathbf{e}^a . \quad (2.30)$$

The bases \mathbf{e}_a and \mathbf{e}^a are said to provide dual bases for the tangent and the cotangent spaces at p .

Tensors of any Rank

Just as a dual vector is a linear map from vectors to the reals, a general tensor can be defined the same way. Let

$$\Pi_s^r \equiv T_p^* M \times \cdots \times T_p^* M \times T_p M \times \cdots \times T_p M \quad (2.31)$$

represent the Cartesian product of r cotangent spaces and s tangent spaces at some point p on the manifold M , i.e. the space of ordered sets of r one-forms and s tangent vectors: $(\omega^1, \dots, \omega^r, X_1, \dots, X_s)$. We now consider a multilinear mapping T of Π_s^r

to the reals. This mapping provides an association of any ordered set of r one-forms and s tangent vectors to the real number, $T(\omega^1, \dots, \omega^r, X_1, \dots, X_s)$. The condition that the map is multilinear requires that

$$\begin{aligned} T(\omega^1, \dots, \omega^r, \alpha X + \beta Y, X_2, \dots, X_s) = \\ \alpha T(\omega^1, \dots, \omega^r, X, X_2, \dots, X_s) + \beta T(\omega^1, \dots, \omega^r, Y, X_2, \dots, X_s) \end{aligned} \quad (2.32)$$

for all $\alpha, \beta \in \mathbf{R}$ and $X, Y \in T_p M$, and for similar replacements of all other forms and vectors. A multilinear mapping so defined is said to be a **tensor of type** (r, s) . Linear combinations of tensors of type (r, s) are defined by the rule

$$\begin{aligned} (\alpha T + \beta S)(\omega^1, \dots, \omega^r, X_1, \dots, X_s) = \\ \alpha T(\omega^1, \dots, \omega^r, X_1, \dots, X_s) + \beta S(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \end{aligned} \quad (2.33)$$

for all real numbers α and β and $X_j \in T_p M$ and $\omega^j \in T_p^* M$. By these rules, tensors of a given type (r, s) span a linear vector-space of dimension n^{r+s} .

We now define a new product known as the **tensor product** denoted by \otimes . If T is a (k, l) tensor and S a (m, n) tensor, we define a $(k+m, l+n)$ tensor by $T \otimes S$

$$\begin{aligned} T \otimes S (\omega^1, \dots, \omega^k, \dots, \omega^{k+m}, X_1, \dots, X_l, \dots, X_{l+n}) = \\ T(\omega^1, \dots, \omega^k, X_1, \dots, X_l) \times S(\omega^{k+1}, \dots, \omega^{k+m}, X_{l+1}, \dots, X_{l+n}). \end{aligned} \quad (2.34)$$

The space of all tensors of type (r, s) is called the space of **tensor products**, denoted by $(r$ tangent spaces and s cotangent spaces)

$$\Pi_s^r(p) = T_p \otimes \cdots \otimes T_p \otimes T_p^* \otimes \cdots \otimes T_p^*. \quad (2.35)$$

A basis of the tensor product space is then provided by the n^{r+s} special mappings

$$\mathbf{e}_{i_1 \dots i_r}{}^{j_1 \dots j_s} = \omega^1_{i_1} \cdots \omega^r_{i_r} X_1^{j_1} \cdots X_s^{j_s}. \quad (2.36)$$

One generally writes this as a tensor product of basis elements

$$\mathbf{e}_{i_1 \dots i_r}{}^{j_1 \dots j_s} = \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_s}. \quad (2.37)$$

In this notation, the tensor product

$$Y_1 \otimes \cdots \otimes Y_r \otimes \Omega^1 \otimes \cdots \otimes \Omega^s \quad (2.38)$$

of r tangent vectors and s one-forms is that element of Π_s^r which maps $(\omega^1, \dots, \omega^r, X_1, \dots, X_s)$ to the number

$$\langle \omega^1, Y_1 \rangle \cdots \langle \omega^r, Y_r \rangle \langle \Omega^1, X_1 \rangle \cdots \langle \Omega^s, X_s \rangle. \quad (2.39)$$

If instead of the dual basis \mathbf{e}_i and \mathbf{e}^j , we choose a different dual basis $\mathbf{e}_{i'}$ and $\mathbf{e}^{j'}$, then it follows that the components of T , relative to the new basis

$$\mathbf{e}_{i'_1} \otimes \cdots \otimes \mathbf{e}_{j'_r} \otimes \mathbf{e}^{j'_1} \otimes \cdots \otimes \mathbf{e}^{j'_s} \quad (2.40)$$

are given by

$$T^{i'_1 \cdots i'_r}_{j'_1 \cdots j'_s} = \phi_{i_1}^{i'_1} \cdots \phi_{i_r}^{i'_r} \phi_{j'_1}^{j_1} \cdots \phi_{j'_s}^{j_s} T^{i_1 \cdots i_r}_{j_1 \cdots j_s}. \quad (2.41)$$

A tensor T of type (0,2) is said to be **symmetric** or **antisymmetric** if

$$T(X, Y) = T(Y, X) \quad \text{or} \quad T(X, Y) = -T(Y, X) \quad (2.42)$$

for all $X, Y \in T_p M$. This property can be extended to higher rank tensors.

The **components of a tensor** in a coordinate basis can be obtained by acting the tensor on the natural one-forms and vectors, here given in a 4D spacetime,

$$T^{\mu_1 \cdots \mu_k}_{v_1 \cdots v_l} = T(dx^{\mu_1}, \dots, dx^{\mu_k}, \partial_{v_1}, \dots, \partial_{v_l}). \quad (2.43)$$

This is equivalent to the expansion

$$T = T^{\mu_1 \cdots \mu_k}_{v_1 \cdots v_l} \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes dx^{v_1} \otimes \cdots \otimes dx^{v_l}. \quad (2.44)$$

The transformation law for general tensors follows then the same pattern of replacing the Lorentz transformation matrix used in flat space with a matrix representing more general coordinate transformations

$$T^{\mu'_1 \cdots \mu'_k}_{v'_1 \cdots v'_l} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{v'_1}}{\partial x^{v_1}} \cdots \frac{\partial x^{v'_l}}{\partial x^{v_l}} T^{\mu_1 \cdots \mu_k}_{v_1 \cdots v_l}. \quad (2.45)$$

Vector Fields and Tensor Fields

If we assign to every point p of a differentiable manifold M a tangent vector $X_p \in T_p M$, then we call the assignment $X : p \rightarrow X_p$ a vector field on M . If (x^1, \dots, x^n) are local coordinates in an open set $U \subset M$, then for every point $p \in U$, X_p has a unique representation of the form

$$X_p = \xi^i(p) \left(\frac{\partial}{\partial x^i} \right)_p. \quad (2.46)$$

The n functions $\xi^i(p)$ defined on U are the components of X with respect to the local coordinate basis.

We denote the set of all C^∞ vector fields on M as $\mathcal{X}(M)$. The four-velocity U on a spacetime is a simple example of a vector field, locally expressed as follows

$$U = u^\mu(s) \partial_\mu, \quad u^\mu = \frac{dx^\mu}{ds}. \quad (2.47)$$

Similarly to vector fields, we also can define tensor fields by assigning to every point p of a differentiable manifold a tensor T . The energy-momentum tensor for pressureless particles with density ϱ is an example

$$T = \varrho U \otimes U = \varrho u^\mu u^\nu \partial_\mu \otimes \partial_\nu. \quad (2.48)$$

Vector Bundles

The tangent space TM is a special example of a vector bundle [11]. A **vector bundle** is a geometrical construct, where to every point of a topological space (here a manifold) we attach a vector space in a compatible way, so that all these vector spaces, glued together, form another topological space. A **real vector bundle** is defined as follows:

- there are two topological spaces M (base space) and E (called total space);
- there exists a continuous map $\pi : E \rightarrow M$ (called projection);
- for every $p \in M$, the fiber $\pi^{-1}(p)$ has the structure of a real vector space; more exactly: for every point in M there is an open neighborhood U and a homeomorphism $\phi : U \times R^n \rightarrow \pi^{-1}(U)$ such that for every point $p \in U$:
- $\pi \circ \phi(p, v) = p$ for all vectors $v \in R^n$;
- the map $v \rightarrow \phi(p, v)$ yields an isomorphism between the vector spaces R^n and $\pi^{-1}(p)$.

The open neighborhood U together with the homeomorphism ϕ is called a local trivialization of the bundle. A vector bundle is called trivial, if there is a global trivialization. TM for M_4 as Minkowski space is a trivial bundle, $TM = M_4 \times M_4$.

Given a vector bundle $\pi : E \rightarrow M$ and an open subset U of M , we can consider **sections** of π on U , i.e. continuous functions $\psi : U \rightarrow E$ with $\pi \circ \psi = id_U$. **Essentially, a section assigns to every point in U a vector from the attached vector space.** Sections of the tangent bundle are nothing but vector fields on the manifold. Let Ψ^A ($A = 1, 2, 3$) be a triplet of Dirac spinors on Minkowski space, then this defines a section into a 3D complex vector space (e.g. the color space of strong interaction).

2.3.2 The Metric Field and Pseudo-Riemannian Manifolds

A **pseudo-Riemannian metric** on a differentiable manifold M is defined as a symmetric $(0, 2)$ tensor field g which has the properties

- $g(X, Y) = g(Y, X)$ for all $X, Y \in \mathfrak{X}(M)$;
- for every $p \in M$, g_p is nondegenerate bilinear form on $T_p M$. This means that $g_p(X, Y) = 0$ for all $X \in T_p M$ if and only if $Y = 0$.

g is called a proper Riemann metric if g_p is positive definite at every point p .

Spacetime is a four-dimensional manifold M with its tangent space $T_p M$ equivalent to Minkowski spacetime, $T_p M = M_4$. A spacetime requires then the existence of a global symmetric $(0, 2)$ -tensor field g which is nondegenerate, i.e. the determinant $g = |g_{\mu\nu}|$ does not vanish, when g is expressed with respect to the coordinate basis. This allows us to define the inverse metric elements $g^{\mu\nu}$ via²

² In 4D spacetimes tensor components given with respect to the coordinate basis are labelled by Greek letters, $\alpha, \beta, \gamma = 0, 1, 2, 3$, their spatial components by $i, k = 1, 2, 3$ and their frame components by $a, b, c = 0, 1, 2, 3$.

$$g^{\mu\nu} g_{\nu\alpha} = \delta_\alpha^\mu. \quad (2.49)$$

As in special relativity, this metric tensor field allows us to raise and lower tensor indices.

Similar to Minkowski space, the metric g enables the nonzero vectors at a point p to be divided into three classes: a nonzero vector X in $T_p M$ is said to be **time-like**, **space-like** or **null** according to whether $g(X, X)$ is positive, negative or zero respectively. The metric therefore determines the causal structure of the spacetime.

A metric tensor (or pseudometric in spacetime) allows us to define orthonormal frames \mathbf{e}_a as special basis elements of the tangent space. With respect to such a **vierbein** basis (also called a tetrad or general observer) \mathbf{e}_a ($a = 0, \dots, 3$) the metric assumes its Minkowski expression, i.e.

$$g(\mathbf{e}_a, \mathbf{e}_b) = \eta_{ab}. \quad (2.50)$$

In this sense, the metric has the same signature as the Minkowski metric η .

In the discussion of path lengths in special relativity we introduced the line element as $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$. We generalize this notion to curved spaces to write down a line element in terms of the components of the metric tensor field g given with its components $g_{\mu\nu}$

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (2.51)$$

In a way the terms metric and line element are used in the literature interchangeably. This is perfect if you think in terms of the line element on the unit sphere S^2

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2. \quad (2.52)$$

This is completely consistent with the interpretation of ds as an infinitesimal length. What does the line element (2.51) mean exactly? In the theory of parametric surfaces and in a Riemannian manifold, the metric tensor allows us in fact to **measure arclengths for curves** $x^j(\lambda)$

$$ds = \sqrt{g_{ik} \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}} d\lambda, \quad (2.53)$$

where g_{ik} are the components of the metric tensor g with respect to the natural basis and $dx^j/d\lambda$ gives the tangent vector to the curve $x^j(\lambda)$ (see (2.14)). The total arclength s of a curve between two points is then given by the integral

$$s = \int_1^2 \sqrt{g_{ik} \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}} d\lambda. \quad (2.54)$$

This arclength is of course a scalar, since it should not depend on the chosen coordinate system.

In a pseudo-Riemannian manifold with a metric tensor g of signature $(-+++)$, this interpretation is only meaningful for **time-like** curves, where the tangent vector

U satisfies $g(U, U) < 0$. As in Minkowski space, the arclength is now the **proper time**, $d\tau = ds/c$, measured by a clock moving along the worldline $x^\mu(\lambda)$ (e.g. a clock in a satellite orbiting the Earth)

$$\tau_{ab} = \int_a^b d\tau = \int_a^b \frac{d\tau}{d\lambda} d\lambda = \frac{1}{c} \int_a^b \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda. \quad (2.55)$$

2.3.3 The Calculus of Forms on Lorentzian Manifolds

A particularly important class of tensors of type $(0, s)$ is the class of totally antisymmetric tensors, i.e. covariant tensors which are antisymmetric in every pair of their arguments. Tensors of this type can be constructed out of a general tensor of type $(0, s)$ by applying to it the alternating operator A defined as

$$AT(X_1, \dots, X_s) \equiv \frac{1}{s!} \sum_{j_1, \dots, j_s} \text{sgn}(j_1, \dots, j_s) T(X_{j_1}, \dots, X_{j_s}), \quad (2.56)$$

where the summation is extended over all $s!$ permutations of the s integers $(1, \dots, s)$ and $\text{sgn}(j_1, \dots, j_s) = \pm 1$ according to as (j_1, \dots, j_s) is an even or odd permutation of $(1, \dots, s)$. If $s > \dim(M)$, then the result is zero, i.e. **there can be no totally antisymmetric tensor of type $(0, s)$ for $s > \dim(M)$** .

Totally antisymmetric tensors of type $(0, p)$ are called **p -forms**. These differential forms are extremely helpful concepts in direct calculation. A zero-form is a scalar function. The one-forms Θ^a defined above are the basis elements of the cotangent space, its components are the components of covariant vectors. A general one-form A can always be written as $A = A_\mu dx^\mu = A_a \Theta^a$. The vector potential of classical electrodynamics is the standard example. A new operation introduced when one works with forms is called the **wedge product**. Given any p -form Ω and a q -form Π , we form the wedge product by the rule

$$\Omega \wedge \Pi = A(\Omega \otimes \Pi). \quad (2.57)$$

This defines now a $(p + q)$ -form if $p + q \leq \dim(M)$. An example of a p -form is

$$A = \frac{1}{p!} A_{\mu\nu\dots\varrho} dx^\mu \wedge dx^\nu \wedge \dots \wedge dx^\varrho, \quad (2.58)$$

where $A_{\mu\nu\dots\varrho}$ is a completely antisymmetric tensor with p indices. In fact, the set of p -forms in a n -dimensional manifold is a vector space Λ_p of dimension $n! / p!(n-p)!$

Table 2.1. Number of linearly independent p -forms for $D = 3$ and $D = 4$

Forms	0	1	2	3	4
dim = 3	1	3	3	1	—
dim = 4	1	4	6	4	1

(see Table 2.1). In four dimensions we have one zero-form, four one-forms (basis in the cotangent space), six two-forms (e.g. the Faraday tensor), four three-forms (currents) and only one four-form (volume-form).

Exterior Derivative

The **exterior derivative** d maps a p -form into a $(p + 1)$ -form, e.g. a one-form

$$d\Theta = d(\Theta_\mu dx^\mu) = \Theta_{\mu,\nu} dx^\nu \wedge dx^\mu = \frac{1}{2} (\Theta_{\mu,\nu} - \Theta_{\nu,\mu}) dx^\nu \wedge dx^\mu. \quad (2.59)$$

In general, for a p -form A given by

$$A = A_{\mu\nu\dots\varrho} dx^\mu \wedge dx^\nu \wedge \dots \wedge dx^\varrho \quad (2.60)$$

the exterior derivative is given by its local expression

$$\begin{aligned} dA &= dA_{\mu\nu\dots\varrho} \wedge dx^\mu \wedge dx^\nu \wedge \dots \wedge dx^\varrho \\ &= \frac{\partial A_{\mu\nu\dots\varrho}}{\partial x^\lambda} dx^\lambda \wedge dx^\mu \wedge dx^\nu \wedge \dots \wedge dx^\varrho. \end{aligned} \quad (2.61)$$

With this explicit definition, one can show

$$d(A \wedge B) = dA \wedge B + (-1)^p A \wedge dB \quad (2.62)$$

$$d(dA) = 0 \quad (2.63)$$

for p -form A and a q -form B .

Inner Product

One can also define an antiderivation i which makes a $(p - 1)$ -form out of a p -form defined as

$$(i_V \omega)(V_1, \dots, V_{p-1}) = \omega(V, V_1, \dots, V_{p-1}), \quad (2.64)$$

i.e. just by contraction with the first index. With the Faraday tensor F we can, for example, build the one-form $E = i_V F$, in components $E_\nu = V^\mu F_{\mu\nu}$. This operation is called the *inner product* of V with ω .

Lie Derivative

Applying both operations, the inner product and the exterior derivative, leaves the degree of a p -form invariant

$$L_X = d \circ i_X + i_X \circ d \quad (2.65)$$

is equivalent to the Lie derivative on p -forms. The **Lie derivative** L is given by its action on functions

$$L_X f = X.f = df(X), \quad (2.66)$$

its action on vector fields

$$L_X Y = [X, Y] \quad (2.67)$$

and the Leibniz rule for the compatibility with higher rank tensors

$$L_X(S \otimes T) = L_X S \otimes T + S \otimes L_X T. \quad (2.68)$$

From the last property, we can derive for a one-form ω and a vector field Y

$$L_X(\omega \otimes Y) = L_X(\omega(Y)) = (L_X\omega) \otimes Y + \omega \otimes (L_X Y). \quad (2.69)$$

Writing out in components, we have

$$X^\mu (\omega_\alpha X^\alpha)_{,\mu} = (L_X\omega)_\mu Y^\mu + \omega_\mu (L_X Y)^\mu, \quad (2.70)$$

or making use of the Lie bracket

$$\begin{aligned} (L_X\omega)_\mu Y^\mu &= X^\alpha (\omega_{\mu,\alpha} Y^\mu + \omega_\mu Y^\mu_{,\alpha}) - \omega_\mu (X^\alpha Y^\mu_{,\alpha} - Y^\alpha X^\mu_{,\alpha}) \\ &= (X^\alpha \omega_{\mu,\alpha} + \omega_\alpha X^\alpha_{,\mu}) Y^\mu. \end{aligned} \quad (2.71)$$

Since this last equation is valid for any vector field Y , we conclude

$$(L_X\omega)_\mu = X^\alpha \omega_{\mu,\alpha} + \omega_\alpha X^\alpha_{,\mu}. \quad (2.72)$$

Volume Element on a Manifold

Each manifold of dimension n carries one n -form $dx^1 \wedge \cdots \wedge dx^n$. With this n -form we can define a volume form

$$\eta \equiv \sqrt{|g|} dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n. \quad (2.73)$$

Due to the transformation of the determinant of the metric, this form is independent of the chosen coordinate system and corresponds to the classical volume element. We can evaluate this form on the vector field $\partial/\partial x^1, \dots, \partial/\partial x^n$ which yields the volume of the parallelepiped spanned by the basis vectors $\partial/\partial x^j$. η is a totally antisymmetric tensor

$$\eta = \frac{1}{n!} \eta_{i_1 \dots i_n} dx^{i_1} \wedge \cdots \wedge dx^{i_n} \quad (2.74)$$

with local components

$$\eta_{i_1 \dots i_n} = \sqrt{|g|} \epsilon_{i_1 \dots i_n}. \quad (2.75)$$

2.4 Affine Connection and Covariant Derivative

A general Lorentzian manifold (M, g) disposes of structures which are not available in Minkowski space. Differentiable manifolds in general can be given an **affine connection** or, in more geometrical terms, it can be defined what means the parallel transport of a tangent vector from one point to another. With help of this affine connection, one can define a covariant derivative, where one has to subtract tensor fields at different spacetime points, which does not even make sense in a manifold without any affine connection, because the tangent vectors live in different tangent spaces. Rather, one has to “parallel transport” the tensor quantities from one point of the manifold to another.

2.4.1 Affine Connection

An **affine connection** is given by a **bilinear mapping** $\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$, which maps an arbitrary vector field Y into a vector field $\nabla_X Y$ and satisfies

- $\nabla_X Y$ is linear in the argument X , i.e.

$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z \quad (2.76)$$

where f and g are arbitrary functions on M ;

- $\nabla_X Y$ is linear in the argument Y , i.e.

$$\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z \quad (2.77)$$

- for any function f we have

$$\nabla_X f = X.f, \quad (2.78)$$

as well as

$$\nabla_X(fY) = (\nabla_X f)Y + f\nabla_X Y. \quad (2.79)$$

2.4.2 Covariant Derivative of Vector Fields

With the action of ∇_X on vector fields Y we now define the **covariant derivative** ∇Y of Y as a tensor field of type $(1,1)$ which maps the contravariant vector field X to $\nabla_X Y$

$$\nabla Y(X) = \langle \nabla Y, X \rangle = \nabla_X Y. \quad (2.80)$$

To clarify what the assignment of a connection precisely means, it will be useful to rewrite $\nabla_X Y$ relative to some dual bases

$$\nabla_X Y = \nabla_X(Y^j \mathbf{e}_j) = (X.Y^j)\mathbf{e}_j + Y^j \nabla_X \mathbf{e}_j. \quad (2.81)$$

Since $\nabla_X \mathbf{e}_j$ is a tensor field of type (1,1), we must have a representation in the given basis

$$\boxed{\nabla_X \mathbf{e}_j = \omega_j^k(X) \mathbf{e}_k}, \quad (2.82)$$

where ω_j^k are one-forms. Accordingly, we may write

$$\nabla_X Y = (X.Y^j) \mathbf{e}_j + Y^j \omega_j^k(X) \mathbf{e}_k. \quad (2.83)$$

Let

$$\omega_j^k(\mathbf{e}_m) \equiv \omega_{mj}^k \quad (2.84)$$

be the coefficients of the expansion of ω_j^k .

From here we conclude that a connection ∇ is specified by the n^2 one-forms ω_j^k (called **connection one-forms**), i.e. by n^3 scalar fields ω_{mj}^k , where $n = \text{Dim}(M)$. In a 4D spacetime, a connection is given by 64 coefficients.

We now return to formula (2.83) and write it in the form

$$\boxed{\nabla_X Y = \left[X.Y^j + \omega_k^j(X)Y^k \right] \mathbf{e}_j}. \quad (2.85)$$

From this we infer that

$$(\nabla_X Y)^j = X.Y^j + \omega_k^j(X)Y^k. \quad (2.86)$$

In a local basis this gives

$$(\nabla_{\partial_m} Y)^j = \partial_m Y^j + Y^k \Gamma_{mk}^j, \quad (2.87)$$

where we have introduced the **local connection coefficients**

$$\boxed{\Gamma_{mk}^j \equiv \omega_k^j(\partial_m)}. \quad (2.88)$$

This is then often written in the standard formula

$$\boxed{Y_{;m}^j = Y_{,m}^j + \Gamma_{mk}^j Y^k} \quad (2.89)$$

and called the **covariant derivative of the vector field** Y given in a coordinate basis.

2.4.3 Covariant Derivative for Tensor Fields

The definition of the covariant derivative of vector fields can be extended to tensor fields, by requiring that the operation of ∇ satisfies the Leibniz rule when acting on tensor products

$$\nabla(S \otimes T) = \nabla S \otimes T + S \otimes \nabla T. \quad (2.90)$$

Covariant Derivative of One-Forms:

We apply this rule to the special tensor product between a one-form α and a vector field Y

$$\nabla_X(\alpha \otimes Y) = (\nabla_X \alpha) \otimes Y + \alpha \otimes (\nabla_X Y). \quad (2.91)$$

With contraction we can write this as

$$\nabla_X(\alpha(Y)) = (\nabla_X \alpha)(Y) + \alpha(\nabla_X Y), \quad (2.92)$$

or as, since $\alpha(Y) = \alpha_j Y^j$ is a scalar,

$$\boxed{(\nabla_X \alpha)(Y) = X.\alpha(Y) - \alpha(\nabla_X Y).} \quad (2.93)$$

We can write this in components

$$\begin{aligned} (\nabla_X \alpha)_j Y^j &= X^k (\alpha_j Y^j)_{,k} - \alpha_k (\nabla_X Y)^k \\ &= X^k \alpha_{j,k} Y^j + X^k \alpha_j Y^j_{,k} - \alpha_k X^j Y^k_{,j} - \alpha_k \omega^k_j Y^j \\ &= X^k \alpha_{j,k} Y^j - \alpha_k \omega^k_j Y^j. \end{aligned} \quad (2.94)$$

This leads us to the classical expression

$$\boxed{(\nabla_X \alpha)_j = X^k \alpha_{j;k} = X^k \alpha_{j,k} - \alpha_k \omega^k_j(X).} \quad (2.95)$$

In a local coordinate basis, we use

$$\nabla_X(\partial_i) = X^k \nabla_{\partial_k}(\partial_i) = X^k \Gamma_{ki}^l \partial_l. \quad (2.96)$$

In addition we have according to (2.93)

$$(\nabla_X dx^j)(\partial_i) = X \cdot \langle dx^j, \partial_i \rangle - \langle dx^j, \nabla_X \partial_i \rangle = -X^k \Gamma_{ki}^j, \quad (2.97)$$

or

$$\boxed{\nabla_X dx^j = -X^k \Gamma_{ki}^j dx^i.} \quad (2.98)$$

The above formula can be generalized to any tensor product of one-forms ω^i and vector fields Y_i and a tensor $S \in \Pi_q^p$

$$\begin{aligned} \nabla_X(\omega^1 \otimes \cdots \otimes \omega^p \otimes Y_1 \otimes \cdots \otimes Y_q \otimes S) &= \\ \nabla_X \omega^1 \otimes \cdots \otimes \omega^p \otimes Y_1 \otimes \cdots \otimes Y_q \otimes S &+ \cdots + \\ \omega^1 \otimes \cdots \otimes \omega^p \otimes \nabla_X Y_1 \otimes Y_2 \otimes \cdots \otimes Y_q \otimes S &+ \cdots + \\ \omega^1 \otimes \cdots \otimes \omega^p \otimes Y_1 \otimes \cdots \otimes Y_q \otimes \nabla_X S. & \end{aligned} \quad (2.99)$$

After a complete contraction with the corresponding components in S we get

$$\begin{aligned}\nabla_X[S(\omega^1, \dots, \omega^p, Y_1, \dots, Y_q)] = \\ S(\nabla_X \omega^1, \dots, \omega^p, Y_1, \dots) + \dots + S(\omega^1, \dots, \omega^p, Y_1, \dots, \nabla_X Y_q) \\ + (\nabla_X S)(\omega^1, \dots, Y_q).\end{aligned}\quad (2.100)$$

From this we find the following expression for $\nabla_X S$

$$\begin{aligned}(\nabla_X S)(\omega^1, \dots, \omega^p, Y_1, \dots, Y_q) = \\ X.S(\omega^1, \dots, \omega^p, Y_1, \dots, Y_q) - S(\nabla_X \omega^1, \omega^2, \dots, Y_q) - \dots \\ - S(\omega^1, \dots, \omega^p, Y_1, \dots, \nabla_X Y_q).\end{aligned}\quad (2.101)$$

Finally, we give the local expression for this covariant derivative of a tensor field $S \in \Pi_q^p(M)$

$$S = S^{i_1 \dots i_p}_{j_1 \dots j_q} \partial_{i_1} \otimes \dots \otimes \partial_{i_p} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_q}. \quad (2.102)$$

With the above expressions for $\nabla_X(\partial_i)$ and $\nabla_X dx^j$ we obtain an expression for the components of ∇S denoted by $S^{i_1 \dots i_p}_{j_1 \dots j_q; k} \equiv \nabla_k S^{i_1 \dots i_p}_{j_1 \dots j_q}$

$S^{i_1 \dots i_p}_{j_1 \dots j_q; k} = S^{i_1 \dots i_p}_{j_1 \dots j_q, k} + \Gamma_{kl}^{i_1} S^{li_2 \dots i_p}_{j_1 \dots j_q} + \dots - \Gamma_{kj_1}^l S^{i_1 \dots i_p}_{lj_2 \dots j_q} - \dots$

(2.103)

The covariant derivatives of vector fields $X = \xi^i(x)\partial_i$ and one-forms $\alpha = \alpha_i(x)dx^i$ are then special cases

$$\xi^i_{,k} = \xi^i_{,k} + \Gamma_{km}^i \xi^m \quad (2.104)$$

$$\alpha_{i;k} = \alpha_{i,k} - \alpha_m \Gamma_{ki}^m. \quad (2.105)$$

2.4.4 Parallel Transport and Metric Connection

Having set up the machinery of connections, the first thing we will do is discuss parallel transport. Recall that in flat space it was unnecessary to be very careful about the fact that vectors were elements of tangent spaces defined at individual points; it is actually very natural to compare vectors at different points (where by “compare” we mean add, subtract, take the dot product, etc.). The reason why it is natural is because it makes sense, in flat space, to “move a vector from one point to another while keeping it constant.” Then, once we get the vector from one point to another, we can do the usual operations allowed in a vector space.

The concept of moving a vector along a path, keeping it constant all the while, is known as **parallel transport**. As we shall see, parallel transport is defined whenever we have a connection; the intuitive manipulation of vectors in flat space makes implicit use of the Christoffel connection on this space. The crucial difference between flat and curved spaces is that, in a curved space, the result of parallel transporting a vector from one point to another will depend on the path taken between the points. Without yet assembling the complete mechanism of parallel transport,

we can use our intuition about the two-sphere to see that this is the case. Start with a vector on the equator, pointing along a line of constant longitude. Parallel transport it up to the north pole along a line of longitude in the obvious way. Then take the original vector, parallel transport it along the equator by an angle θ , and then move it up to the north pole as before. It is clear that the vector, parallel transported along two paths, arrived at the same destination with two different values (rotated by θ).

We simply must learn to live with the fact that two vectors can only be compared in a natural way if they are elements of the same tangent space. For example, two particles passing by each other have a well-defined relative velocity (which cannot be greater than the speed of light). But two particles at different points on a curved manifold do not have any well-defined notion of relative velocity – the concept simply makes no sense. Of course, in certain special situations it is still useful to talk as if it did make sense, but it is necessary to understand that occasional usefulness is not a substitute for rigorous definition. In cosmology, for example, the light from distant galaxies is redshifted with respect to the frequencies we would observe from a nearby stationary source. Since this phenomenon bears such a close resemblance to the conventional Doppler effect due to relative motion, it is very tempting to say that the galaxies are “receding away from us” at a speed defined by their redshift. At a rigorous level this is nonsense – the galaxies are not receding, since the notion of their velocity with respect to us is not well-defined. What is actually happening is that the metric of spacetime between us and the galaxies has changed (the Universe has expanded) along the path of the photon from here to there, leading to an increase in the wavelength of the light. As an example of how you can go wrong, naive application of the Doppler formula to the redshift of galaxies implies that some of them are receding faster than light, in apparent contradiction with relativity. The resolution of this apparent paradox is simply that the very notion of their recession should not be taken literally.

Parallel transport is supposed to be the curved-space generalization of the concept of “keeping the vector constant” as we move it along a path; similarly for a tensor of arbitrary rank. Given a curve $x^\mu(\lambda)$, the requirement of constancy of a tensor T along this curve in flat space is simply

$$\frac{dT}{d\lambda} \equiv u^\mu \frac{\partial T}{\partial x^\mu} = 0, \quad (2.106)$$

where $u^\mu = dx^\mu/d\lambda$ is the four-velocity. We therefore define the covariant derivative along the path to be given by an operator

$$\frac{D}{d\lambda} \equiv u^\mu \nabla_\mu. \quad (2.107)$$

This is a well-defined tensor equation, since both the tangent vector $dx^\mu/d\lambda$ and the covariant derivative ∇T are tensors. This is known as the **equation of parallel transport**. For a vector it takes the form

$$\frac{d}{d\lambda} V^\mu + u^\alpha \Gamma_{\alpha\beta}^\mu V^\beta = 0. \quad (2.108)$$

We can look at the parallel transport equation as a first-order differential equation defining an initial-value problem: given a tensor at some point along the path, there will be a unique continuation of the tensor to other points along the path such that the continuation solves (2.108). We say that such a tensor is **parallel transported**.

As an aside, an especially interesting example of the parallel propagator occurs when the path is a loop, starting and ending at the same point. Then if the connection is metric-compatible, the resulting matrix will just be a Lorentz transformation on the tangent space at the point. This transformation is known as the **holonomy** of the loop. If you know the holonomy of every possible loop, that turns out to be equivalent to knowing the metric. This fact has led Ashtekar and his collaborators to examine general relativity in the **loop representation**, where the fundamental variables are holonomies rather than the explicit metric. They have made some progress towards quantizing the theory in this approach, although the jury is still out about how much further progress can be made [12].

2.4.5 Metric Connection

The notion of parallel transport is obviously dependent on the connection, and different connections lead to different answers. We say:

An affine connection is a metric connection if parallel transport along any smooth curve γ in M preserves the inner product in the sense that for autoparallel fields $X(\lambda)$ and $Y(\lambda)$ given along γ , $g(X(\lambda), Y(\lambda))$ is independent of λ .

If the connection is **metric-compatible**, the metric is always parallel transported with respect to it

$$\frac{D}{d\lambda} g = 0 \quad , \quad \nabla g = 0. \quad (2.109)$$

It follows that the inner product of two parallel-transported vectors is preserved. That is, if X and Y are parallel-transported along a curve $x^\mu(\lambda)$, we have

$$\begin{aligned} \frac{D}{d\lambda} g(X, Y) &= \left(\frac{D}{d\lambda} g_{\mu\nu} \right) X^\mu Y^\nu \\ &+ g_{\mu\nu} X^\mu \frac{D}{d\lambda} Y^\nu + g_{\mu\nu} X^\mu \frac{D}{d\lambda} Y^\nu = 0. \end{aligned} \quad (2.110)$$

This means that *parallel transport with respect to a metric-compatible connection preserves the norm of vectors*, the sense of orthogonality.

As a result, $\nabla g = 0$ is equivalent to the Ricci identity

$$X.g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \quad (2.111)$$

From this we can conclude

For every pseudo-Riemannian manifold (M, g) , there exists a unique affine connection ∇ such that

- ∇ is symmetric (i.e. torsion vanishes);
- ∇ is metric.

∇ is called symmetric if the torsion tensor T vanishes which is defined as (see next section)

$$\boxed{T(X, Y) \equiv \nabla_X Y - \nabla_Y X - [X, Y] = 0} \quad (2.112)$$

for all vector fields X and Y . In local coordinates this just means $\Gamma_{km}^i = \Gamma_{mk}^i$. Since torsion vanishes, the Ricci identity implies

$$X.g(Y, Z) = g(\nabla_Y X, Z) + g([X, Y], Z) + g(Y, \nabla_X Z). \quad (2.113)$$

From cyclic permutations we also obtain

$$Y.g(Z, X) = g(\nabla_Z Y, X) + g([Y, Z], X) + g(Z, \nabla_Y X) \quad (2.114)$$

and

$$Z.g(X, Y) = g(\nabla_X Z, Y) + g([Z, X], Y) + g(X, \nabla_Z Y). \quad (2.115)$$

Taking the linear combination of the second plus the third minus the first relation results in

$$\begin{aligned} 2g(\nabla_Z Y, X) &= -X.g(Y, Z) + Y.g(Z, X) + Z.g(X, Y) \\ &\quad -g([Z, X], Y) - g([Y, Z], X) + g([X, Y], Z). \end{aligned} \quad (2.116)$$

The right-hand side is independent of ∇ . Since g is nondegenerate, the uniqueness follows from this relation.

This unique connection on (M, g) is called the Riemann connection (or Levi-Civita connection).

If we write the above relation in local coordinates, we obtain

$$\begin{aligned} 2g(\nabla_{\partial_i} \partial_j, \partial_k) &= 2\Gamma_{ij}^m g_{mk} = -\partial_k g(\partial_j, \partial_i) + \partial_j g(\partial_i, \partial_k) + \partial_i g(\partial_k, \partial_j) \\ &= -\partial_k g_{ji} + \partial_j g_{ik} + \partial_i g_{kj}, \end{aligned} \quad (2.117)$$

or

$$g_{mk} \Gamma_{ij}^m = \frac{1}{2} (g_{jk,i} + g_{ik,j} - g_{ij,k}). \quad (2.118)$$

This expression is equivalent to the Christoffel symbols

$$\boxed{\Gamma_{ij}^m = \frac{1}{2} g^{mk} (g_{ki,j} + g_{kj,i} - g_{ij,k})}. \quad (2.119)$$

Particle Motion in Lorentzian Manifolds

It is known that on two-dimensional surfaces one can define geodesics as the shortest curves between two points, i.e.

$$\int_1^2 ds = \text{extremum}. \quad (2.120)$$

In a pseudo-Riemannian space where ds^2 can also be zero or even negative, this formulation is no longer valid. Instead of this we can start from the variational principle along a curve $x^\alpha(\lambda)$ parametrized by the parameter λ

$$E(\dot{x}, \dot{x}) = \int_a^b \mathcal{L} d\lambda = \int_a^b \left(\frac{ds}{d\lambda} \right)^2 d\lambda = \int_a^b g_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} d\lambda = \text{extremum}. \quad (2.121)$$

This is known in differential geometry as the energy integral. The function defined as

$$\mathcal{Q}(x_a, x_b) = \frac{1}{2}(s_b - s_a) \int_a^b g_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} d\lambda \quad (2.122)$$

is also called the **world function**. The Euler–Lagrange equation for the extremal problem associated with the integral is

$$\frac{d}{d\lambda} \frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} - \frac{\partial \mathcal{L}}{\partial x^\alpha} = 0. \quad (2.123)$$

With

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = 2g_{\beta\alpha} \dot{x}^\beta \quad (2.124)$$

$$\frac{\partial \mathcal{L}}{\partial x^\alpha} = g_{\mu\nu,\alpha} \dot{x}^\mu \dot{x}^\nu \quad (2.125)$$

the Euler–Lagrange equation reduces to the condition

$$g_{\beta\alpha} \ddot{x}^\beta + (g_{\mu\nu,\alpha} - \frac{1}{2}g_{\mu\nu,\alpha}) \dot{x}^\mu \dot{x}^\nu = 0. \quad (2.126)$$

The second term can be symmetrized so that

$$g_{\beta\alpha} \ddot{x}^\beta + \frac{1}{2}(g_{\mu\nu,\alpha} + g_{\alpha\mu,\nu} - g_{\mu\nu,\alpha}) \dot{x}^\mu \dot{x}^\nu = 0. \quad (2.127)$$

Now we contract with the inverse metric to obtain the equation

$$\ddot{x}^\alpha + \Gamma_{\mu\nu}^\alpha \dot{x}^\mu \dot{x}^\nu = 0. \quad (2.128)$$

The symbols $\Gamma_{\mu\nu}^\alpha = \Gamma_{\nu\mu}^\alpha$ are called **Christoffel symbols** of the metric g defined by (2.119). The above equation determines the geodetic motion of particles in the metric field g . For time-like curves we can make the statement that force-free particles move on a geodesic, parametrized in terms of the arclength s

$$\frac{d^2 x^\alpha}{ds^2} + \Gamma_{\mu\nu}^\alpha \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0.$$

(2.129)

Newtonian Limit

A slowly moving particle has $dx^0/ds \simeq 1$ and we may neglect dx^i/ds ($i = 1, 2, 3$) in comparison to dx^0/ds . In leading order, the geodesic equations for the pseudo-Newtonian metric are given by

$$\frac{d^2x^i}{dt^2} \simeq \frac{d^2x^i}{d\tau^2} = -\Gamma_{\mu\nu}^i \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \simeq -c^2 \Gamma_{00}^i. \quad (2.130)$$

Thus, only the components Γ_{00}^i appear in the equations of motion with

$$\Gamma_{00}^i = -\frac{1}{2} g^{ik} g_{00,k} = \frac{1}{2} g_{00,i} = \frac{\nabla_i \Phi}{c^2}, \quad (2.131)$$

where we have used the pseudo-Newtonian metric and the fact that the gravitational field is static. This equivalence between the local gravitational force Γ_{00}^i and the Newtonian force dictates the Newtonian limit in the metric expansion.

2.4.6 Divergence of Vector Fields

The divergence of a vector field V is given by

$$V_{;j}^j = V_{,j}^j + \Gamma_{jk}^j V^k. \quad (2.132)$$

In this case, the Christoffel symbols can be simplified

$$\Gamma_{jk}^j = \frac{1}{2} g^{jm} (g_{mj,k} + g_{mk,j} - g_{jk,m}) = \frac{1}{2} g^{jm} g_{mj,k}. \quad (2.133)$$

This expression can be simplified if we recall that for any matrix M

$Tr(M^{-1}(x) \partial_l M(x)) = \partial_l \ln \det(M).$

(2.134)

To prove this, we consider the variation in $\ln \det(M)$ owing to variations δx in x

$$\begin{aligned} \delta \ln \det(M) &= \ln \det(M + \delta M) - \ln \det(M) \\ &= \ln \frac{\det(M + \delta M)}{\det(M)} \\ &= \ln \det[M^{-1}(M + \delta M)] \\ &\simeq \ln[1 + Tr(M^{-1} \delta M)] \\ &\simeq Tr[M^{-1} \delta M]. \end{aligned} \quad (2.135)$$

Replacing the matrix M by the metric g_{ik} we find

$$\Gamma_{jk}^j = \frac{1}{2} \partial_k \ln g = \frac{1}{\sqrt{g}} \partial_k \sqrt{g}, \quad (2.136)$$

where $g = \det(g_{ik})$. Therefore, the covariant divergence of vector fields can always be evaluated without referring to Christoffel symbols

$$V^j_{;j} = \frac{1}{\sqrt{g}} \partial_j(\sqrt{g} V^j) . \quad (2.137)$$

This property can also be used to simplify the divergence of a (2,0)-tensor

$$T^{jk}_{;k} = \partial_k T^{jk} + \Gamma^j_{jm} T^{mk} + \Gamma^k_{jm} T^{jm} . \quad (2.138)$$

If T is antisymmetric, the last term drops out

$$A^{jk}_{;k} = \frac{1}{\sqrt{g}} \partial_k(\sqrt{g} A^{jk}) . \quad (2.139)$$

If T is symmetric (e.g. the energy–momentum tensor), then

$$T^{jk}_{;k} = \frac{1}{\sqrt{g}} \partial_k(\sqrt{g} T^{jk}) + \Gamma^j_{km} T^{km} . \quad (2.140)$$

2.5 Curvature of Pseudo-Riemannian Manifolds

The curvature of a manifold is quantified by the Riemann tensor, which is derived from the connection. The idea behind this measure of curvature is that we know what we mean by “flatness” of a connection – the conventional (and usually implicit) Christoffel connection associated with a Euclidean or Minkowskian metric has a number of properties which can be thought of as different manifestations of flatness. These include the fact that parallel transport around a closed loop leaves a vector unchanged, that covariant derivatives of tensors commute, and that initially parallel geodesics remain parallel. As we shall see, the Riemann tensor arises when we study how any of these properties are altered in more general contexts.

A Heuristic Definition

We have already argued, using the two-sphere as an example, that parallel transport of a vector around a closed loop in a curved space will lead to a transformation of the vector. The resulting transformation depends on the total curvature enclosed by the loop; it would be more useful to have a local description of the curvature at each point, which is what the Riemann tensor is supposed to provide. One conventional way to introduce the Riemann tensor, therefore, is to consider parallel transport around an infinitesimal loop. Imagine that we parallel transport a vector V around a closed loop defined by two vectors A and B .

The (infinitesimal) lengths of the sides of the loop are δa and δb , respectively. Now, we know the action of parallel transport is independent of coordinates, so

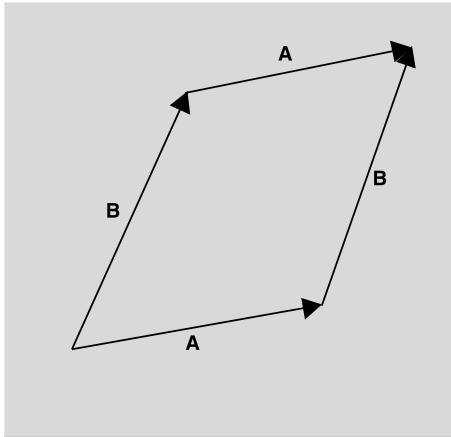


Fig. 2.2. Parallel transport around an infinitesimal loop generated by two vectors A and B defines the Riemann tensor

there should be some tensor which tells us how the vector changes when it comes back to its starting point; it will be a linear transformation on a vector, and therefore involve one upper and one lower index. But it will also depend on the two vectors A and B which define the loop; therefore there should be two additional lower indices to contract with A^ν and B^μ . Furthermore, the tensor should be antisymmetric in these two indices, since interchanging the vectors corresponds to traversing the loop in the opposite direction, and should give the inverse of the original answer. (This is consistent with the fact that the transformation should vanish if A and B are the same vector.) We therefore expect that the expression for the change V experienced by this vector when parallel transported around the loop should be of the form

$$\delta V^\varrho = \delta a \delta b A^\nu B^\mu R_{\sigma\nu\mu}^\varrho V^\sigma. \quad (2.141)$$

R is a $(1, 3)$ -tensor, known as the Riemann tensor.

2.5.1 Mathematical Definition of Torsion and Curvature

Conventionally, the **torsion fields** T are defined as mappings $T : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \rightarrow \mathcal{X}(\mathcal{M})$ on the set of all vector fields on the manifold

$$T(X, Y) \equiv \nabla_X Y - \nabla_Y X - [X, Y]. \quad (2.142)$$

Curvature is defined as a mapping $R : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \rightarrow \mathcal{X}(\mathcal{M})$

$$R(X, Y)Z \equiv \nabla_X (\nabla_Y Z) - \nabla_Y (\nabla_X Z) - \nabla_{[X,Y]} Z. \quad (2.143)$$

They obviously satisfy the antisymmetry conditions

$$T(X, Y) = -T(Y, X), \quad R(X, Y) = -R(Y, X) \quad (2.144)$$

as well as

$$T(fX, gY) = fg T(X, Y) \quad (2.145)$$

$$R(fX, gY)hZ = fgh R(X, Y)Z \quad (2.146)$$

for any functions f, g and h .

The mapping of $\mathcal{X}^*(M) \times \mathcal{X}(M) \times \mathcal{X}(M)$ into the reals defined by

$$(\omega, X, Y) \rightarrow \langle \omega, T(X, Y) \rangle \quad (2.147)$$

is thus a tensor field in $\Pi_2^1(M)$ and is known as the **torsion tensor**. Similarly, a mapping

$$(\omega, Z, X, Y) \rightarrow \langle \omega, R(X, Y)Z \rangle \quad (2.148)$$

is a tensor field in $\Pi_3^1(M)$ and is called **curvature tensor**.

In local coordinates, the components of the torsion tensor are given by

$$T_{ij}^k = \langle dx^k, T(\partial_i, \partial_j) \rangle = \langle dx^k, \nabla_{\partial_i} \partial_j - \nabla_{\partial_j} \partial_i \rangle. \quad (2.149)$$

Please note that the Lie bracket vanishes in local coordinates. Therefore, the torsion tensor gets the expression

$$\boxed{T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k.} \quad (2.150)$$

If torsion vanishes, we have $\Gamma_{ij}^k = \Gamma_{ji}^k$ in every chart. The components of the curvature tensor are then similarly given by

$$\begin{aligned} R^i_{jkl} &= \langle dx^i, R(\partial_k, \partial_l) \partial_j \rangle = \langle dx^i, (\nabla_{\partial_k} \nabla_{\partial_l} - \nabla_{\partial_l} \nabla_{\partial_k}) \partial_j \rangle \\ &= \langle dx^i, \nabla_{\partial_k} (\Gamma_{lj}^s \partial_s) - \nabla_{\partial_l} (\Gamma_{kj}^s \partial_s) \rangle. \end{aligned} \quad (2.151)$$

From this we get the famous expression

$$\boxed{R^i_{jkl} = \Gamma^i_{lj,k} - \Gamma^i_{kj,l} + \Gamma^s_{lj} \Gamma^i_{ks} - \Gamma^s_{kj} \Gamma^i_{ls}.} \quad (2.152)$$

2.5.2 Bianchi Identities for Metric Connection

Torsion and curvature satisfy two identities which are known as **Bianchi identities**:

- **First Bianch identity:**

$$\sum_{\text{cyclic}} \{R(X, Y)Z\} = \sum_{\text{cyclic}} \{T(T(X, Y), Z) + (\nabla_X T)(Y, Z)\}, \quad (2.153)$$

- **Second Bianchi identity:**

$$\sum_{\text{cyclic}} \{(\nabla_X R)(Y, Z) + R(T(X, Y), Z)\} = 0. \quad (2.154)$$

For vanishing torsion, the first Bianchi identity reads as

$$R(X, Y)Z + R(Z, X)Y + R(Y, Z)X = 0 \quad (2.155)$$

or in components, these are four identities

$$\boxed{R^a_{bmn} + R^a_{nbm} + R^a_{mnb} = 0.} \quad (2.156)$$

Geometric indices and internal indices get mixed up. The second Bianchi identity reads as

$$(\nabla_X R)(Y, Z) + (\nabla_Z R)(X, Y) + (\nabla_Y R)(Z, X) = 0, \quad (2.157)$$

or in components

$$\boxed{R^a_{bmn;k} + R^a_{bkm;n} + R^a_{bnk;m} = 0.} \quad (2.158)$$

This reminds us of the homogeneous Maxwell equations for the Faraday tensor F

$$F_{mn,k} + F_{kn,n} + F_{nk,m} = 0. \quad (2.159)$$

In total, the curvature of a metric connection in a Riemannian manifold satisfies the following three relations

$$R^a_{bmn} = -R^a_{bnm} \quad (2.160)$$

$$R_{abmn} = -R_{bamn} \quad (2.161)$$

$$R_{abmn} = R_{mnab}. \quad (2.162)$$

The first relation tells us that the curvature tensor is a two-form, the second relation that the curvature is an element of the Lie algebra of the rotation group, and the third one that tensorial components are exchanged by internal components. This last relation follows from the first Bianchi identity

$$\begin{aligned} R_{jkmn} &= -(R_{jnkm} + R_{jmkn}) \\ &= R_{njkm} + R_{mjnk} \\ &= -(R_{nkmj} + R_{nmjk}) - (R_{mnkj} + R_{mkjn}) \\ &= 2R_{mnjk} + (R_{kmjn} + R_{knmj}) \\ &= 2R_{mnjk} - R_{kjnm} = 2R_{mnjk} - R_{jkmn}. \end{aligned} \quad (2.163)$$

This means that

$$\boxed{R_{jkmn} = R_{mnjk}.} \quad (2.164)$$

Given these relationships between the components of the Riemann tensor, how many independent quantities remain? The Riemann tensor is antisymmetric in the

first two indices, antisymmetric in the last two indices, and symmetric under exchange of these two pairs. We can think of the Riemann tensor as a symmetric matrix $R_{[ab][mn]}$, where the pairs are thought to be individual indices. A $n \times n$ symmetric matrix has $n(n+1)/2$ (these are 10 in 4D) independent components, while the corresponding antisymmetric matrix has $n(n-1)/2$ independent components (these are six in 4D). We therefore have

$$\frac{1}{2} \left[\frac{1}{2}n(n-1) \right] \left[\frac{1}{2}n(n-1)+1 \right] = \frac{1}{8} (n^4 - 2n^3 - 3n^2 - 2n) \quad (2.165)$$

independent components. In addition we have to satisfy the first Bianchi identity

$$R^a_{[bmn]} = 0. \quad (2.166)$$

These are another $n(n-1)(n-2)(n-3)/4!$ constraints. We are left with

$$\frac{1}{8} (n^4 - 2n^3 - 3n^2 - 2n) - \frac{1}{24} (n-1)(n-2)(n-3) = \frac{1}{12} n^2(n^2-1) \quad (2.167)$$

independent components of the Riemann tensor.

In four dimensions, therefore, the **Riemann tensor has 20 independent components**, in two dimensions just one component, and in three dimensions six components. In 4D, the metric tensor only has 10 independent components, but there is more information in the Riemann tensor: 10 components are in the Ricci tensor, the other 10 components are hidden in the Weyl tensor.

2.5.3 Ricci, Weyl and Einstein Tensor

The Riemann tensor R_{ijkl} is antisymmetric in both pairs of indices. By these symmetries, the only nontrivial contraction we can make is that leading to the **Ricci tensor** R_{ij} defined as

$$R_{mn} = R^j_{mjn} = R_{nm} \quad (2.168)$$

and its trace $R = R^m_m$, the **Ricci scalar**. Accordingly, it will be convenient to separate the Riemann tensor into trace-free parts and the Ricci part

$$\begin{aligned} C_{ijkl} &= R_{ijkl} - \frac{1}{n-2} [g_{ik}R_{jl} + g_{jl}R_{ik} - g_{jk}R_{il} - g_{il}R_{jk}] \\ &\quad + \frac{1}{(n-1)(n-2)} R [g_{ik}g_{jl} - g_{il}g_{jk}]. \end{aligned} \quad (2.169)$$

This **Weyl tensor** C_{ijkl} has the same symmetries as the curvature tensor R_{ijkl} . The Weyl tensor is trace-free in the sense that

$$g^{jl}C_{ijkl} = 0. \quad (2.170)$$

The Weyl tensor is only defined, when the manifold carries a metric. The most important property is its conformal invariance against transformations of the type $g \rightarrow \Omega^2(x)g$. **In 4D, 10 components of the curvature tensor are in the Ricci tensor, while the other 10 components are given by the Weyl tensor.**

With the Ricci tensor we also define the **Einstein tensor** G_{ij} by means of

$$\boxed{G_{ij} = R_{ij} - \frac{1}{2} R g_{ij}.} \quad (2.171)$$

This tensor is divergence-free, i.e. $G^i_{j;i} = 0$. This follows from contracting the second Bianchi identity for vanishing torsion

$$R^i_{jkl;m} - g^{ip} R_{jplm;k} + R^i_{jmkl} = 0 \quad (2.172)$$

with respect to i and k

$$R_{jl;m} - g^{kp} R_{jplm;k} - R_{jm;l} = 0. \quad (2.173)$$

Raising j and contracting with m

$$R^j_{l;j} + R^k_{l;k} - R_{,l} = 0 \quad (2.174)$$

or

$$\left(R^j_l - \frac{1}{2} \delta^j_l R \right)_{;j} = 0. \quad (2.175)$$

2.5.4 Cartan's Structure Equations

Since torsion $T(X, Y)$ and curvature $R(X, Y)$ are antisymmetric tensors, they naturally define corresponding two-forms

$$T(X, Y) = \mathcal{T}^a(X, Y) e_a \quad (2.176)$$

$$R(X, Y) e_b = \Omega^a_b(X, Y) e_a. \quad (2.177)$$

The exterior derivatives of the basic one-forms Θ^a and of the connection forms ω satisfy Cartan's structure equations

$$\mathcal{T}^a = d\Theta^a + \omega^a_b \wedge \Theta^b \quad (2.178)$$

$$\Omega^a_b = d\omega^a_b + \omega^a_d \wedge \omega^d_b. \quad (2.179)$$

The wedge operator denotes the exterior products for p -forms. The two-form Ω is the curvature two-form which gives, when expressed locally,

$$\boxed{\Omega^a_b = \frac{1}{2} R^a_{bcd} \Theta^c \wedge \Theta^d} \quad (2.180)$$

the components of the Riemann tensor R^a_{bcd} in orthonormal coordinates. Similarly, we have four torsion two-forms

$$\boxed{\mathcal{T}^a = \frac{1}{2} T^a_{bc} \Theta^b \wedge \Theta^c.} \quad (2.181)$$

For the proof of Cartan's structure equations, we use the above definition of torsion. Written as one-forms, this means

$$\begin{aligned} \mathcal{T}^a(X, Y) e_a &= \nabla_X Y - \nabla_Y X - [X, Y] \\ &= \nabla_X(\Theta^b(Y)e_b) - \nabla_Y(\Theta^b(X)e_b) - \Theta^a([X, Y])e_a \\ &= \{X.\Theta^a(Y) - Y.\Theta^a(X) - \Theta^a([X, Y])\}e_a \\ &\quad + \{\Theta^a(Y)\omega^b_a(X) - \Theta^a(X)\omega^b_a(Y)\}e_a \\ &= d\Theta^a(X, Y)e_a + (\omega^a_b \wedge \Theta^b)(X, Y)e_a. \end{aligned} \quad (2.182)$$

The proof of the second structure equation is similar. Written as a two-form, this means

$$\begin{aligned} \Omega^a_c(X, Y) e_a &= \nabla_X \nabla_Y e_c - \nabla_Y \nabla_X e_c - \omega^b_c([X, Y])e_b \\ &= \nabla_X(\omega^b_c(Y)e_b) - \nabla_Y(\omega^b_c(X)e_b) - \omega^a_c([X, Y])e_a \\ &= \{X.\omega^a_c(Y) - Y.\omega^a_c(X) - \omega^a_c([X, Y])\}e_a \\ &\quad + \{\omega^b_c(Y)\omega^a_b(X) - \omega^b_c(X)\omega^a_b(Y)\}e_a \\ &= d\omega^a_c(X, Y)e_a + (\omega^a_b \wedge \omega^b_c)(X, Y)e_a. \end{aligned} \quad (2.183)$$

Local Expressions

In local coordinates, a metric connection is expressed in terms of the Christoffel symbols

$$\Gamma_{\mu\beta}^\alpha = \frac{1}{2} g^{\alpha\varrho} (g_{\varrho\mu,\beta} + g_{\varrho\beta,\mu} - g_{\mu\beta,\varrho}) \quad (2.184)$$

such that the connection form is given in a local coordinate basis as

$$\omega_\beta^\alpha = \Gamma_{\mu\beta}^\alpha dx^\mu, \quad (2.185)$$

and therefore

$$\begin{aligned} d\omega_\beta^\alpha &= \Gamma_{\mu\beta,v}^\alpha dx^v \wedge dx^\mu \\ &= \frac{1}{2} (\Gamma_{\mu\beta,v}^\alpha - \Gamma_{v\beta,\mu}^\alpha) dx^v \wedge dx^\mu. \end{aligned} \quad (2.186)$$

Also,

$$\begin{aligned} \omega_\varrho^\alpha \wedge \omega_\beta^\varrho &= \Gamma_{v\varrho}^\alpha \Gamma_{\mu\beta}^\varrho dx^v \wedge dx^\mu \\ &= \frac{1}{2} (\Gamma_{v\varrho}^\alpha \Gamma_{\mu\beta}^\varrho - \Gamma_{\mu\varrho}^\alpha \Gamma_{v\beta}^\varrho) dx^v \wedge dx^\mu \end{aligned} \quad (2.187)$$

Accordingly, Cartan's second structure equation is equivalent to the conventional definition of the Riemann tensor in local coordinates

$$R^\alpha_{\beta\mu\nu} = \Gamma^\alpha_{v\beta,\mu} - \Gamma^\alpha_{\mu\beta,v} + \Gamma^\alpha_{\mu\varrho}\Gamma^\varrho_{v\beta} - \Gamma^\alpha_{v\varrho}\Gamma^\varrho_{\mu\beta}. \quad (2.188)$$

An Example: Curvature of a Two-Sphere

A simple example is the two-sphere with the line element given by

$$ds^2 = a^2 (d\theta^2 + \sin^2 \theta d\phi^2), \quad g_{\theta\theta} = a^2, \quad g_{\phi\phi} = a^2 \sin^2 \theta, \quad (2.189)$$

where a is the radius. The nonzero Levi-Civita connection coefficients are

$$\Gamma^\theta_{\phi\phi} = -\sin \theta \cos \theta \quad (2.190)$$

$$\Gamma^\theta_{\theta\phi} = \Gamma^\theta_{\phi\theta} = \cot \theta. \quad (2.191)$$

The Riemann tensor has essentially only one component given by

$$R^\theta_{\phi\theta\phi} = \partial_\theta \Gamma^\theta_{\phi\phi} - \partial_\phi \Gamma^\theta_{\theta\phi} + \Gamma^\theta_{\theta\alpha}\Gamma^\alpha_{\phi\phi} - \Gamma^\theta_{\phi\alpha}\Gamma^\alpha_{\theta\phi} = \sin^2 \theta. \quad (2.192)$$

From this we find the curvature

$$R_{\theta\phi\theta\phi} = g_{\theta\theta} R^\theta_{\phi\theta\phi} = a^2 \sin^2 \theta \quad (2.193)$$

and the corresponding Ricci tensor

$$R_{\theta\theta} = g^{\phi\phi} R_{\phi\theta\phi\theta} = 1 \quad (2.194)$$

$$R_{\theta\phi} = R_{\phi\theta} = 0 \quad (2.195)$$

$$R_{\phi\phi} = g^{\theta\theta} R_{\theta\phi\theta\phi} = \sin^2 \theta. \quad (2.196)$$

The Ricci scalar is found to be given by

$$R = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = \frac{2}{a^2}. \quad (2.197)$$

The Ricci scalar is a constant over this surface. This is a reflection of the fact that the manifold is **maximally symmetric**. The curvature of maximally symmetric manifolds satisfies the relation

$$R_{abmn} = \frac{1}{a^2} [g_{am}g_{bn} - g_{an}g_{bm}]. \quad (2.198)$$

Notice that the Ricci scalar is not only constant for the two-sphere, it is manifestly positive. We say that the sphere is *positively curved*. From the point of view of someone living on a manifold which is embedded in a higher-dimensional Euclidean space, if they are sitting at a point of positive curvature the space curves away from them in the same way in any direction, while in a negatively curved space it curves away in opposite directions. Negatively curved spaces are therefore saddle-like.

Spin Connection of a Two-Sphere

Curvature in the coordinate basis has no invariant meaning. For this we compute the curvature in the orthonormal frame

$$\mathbf{e}_{\hat{\theta}} = \frac{1}{a} \partial_{\theta}, \quad \mathbf{e}_{\hat{\phi}} = \frac{1}{a \sin \theta} \partial_{\phi} \quad (2.199)$$

with the corresponding dual basis

$$\Theta^{\theta} = a d\theta, \quad \Theta^{\phi} = a \sin \theta d\phi. \quad (2.200)$$

With respect to this basis the curvature tensor is given by

$$R_{\hat{\theta}\hat{\phi}\hat{\theta}\hat{\phi}} = e_{\theta}^i e_{\phi}^j e_{\theta}^m e_{\phi}^n R_{ijmn} = \frac{1}{a^2}, \quad (2.201)$$

i.e. here the curvature is one over the square of the radius of the sphere.

The same result can directly be obtained by using the Cartan structure equations

$$d\Theta^{\theta} = 0 \quad (2.202)$$

$$\begin{aligned} d\Theta^{\phi} &= a \cos \theta d\theta \wedge d\phi = \frac{\cos \theta}{a \sin \theta} \Theta^{\theta} \wedge \Theta^{\phi} \\ &= -\frac{\cos \theta}{a \sin \theta} \Theta^{\phi} \wedge \Theta^{\theta} = -\omega_{\theta}^{\phi} \wedge \Theta^{\theta}. \end{aligned} \quad (2.203)$$

The first structure equation tells us that the connection one-form is given by

$$\omega_{\theta}^{\phi} = \frac{\cos \theta}{a \sin \theta} \Theta^{\phi} = \cos \theta d\phi. \quad (2.204)$$

This is called the spin connection of the two-sphere. Since ω_{θ}^{ϕ} is antisymmetric, there is only one such connection form in 2D. Taking the exterior derivative, we arrive at the curvature two-form

$$\begin{aligned} \Omega^{\phi} &= d\omega_{\theta}^{\phi} + \omega_{\theta}^{\phi} \wedge \omega_{\theta}^j = d\omega_{\theta}^{\phi} = -\sin \theta d\theta \wedge d\phi \\ &= -\frac{1}{a^2} \Theta^{\theta} \wedge \Theta^{\phi} = \frac{1}{a^2} \Theta^{\phi} \wedge \Theta^{\theta}. \end{aligned} \quad (2.205)$$

In 2D, the second contribution vanishes (since $\text{SO}(2)$ is Abelian). From this we can directly read off the curvature component

$$\Omega_{\theta}^{\phi} = R_{\hat{\theta}\hat{\phi}\hat{\theta}}^{\hat{\phi}} \Theta^{\phi} \wedge \Theta^{\theta} \quad (2.206)$$

with

$$R_{\hat{\theta}\hat{\phi}\hat{\theta}}^{\hat{\phi}} = R_{\hat{\phi}\hat{\theta}\hat{\phi}\hat{\theta}} = \frac{1}{a^2} = R_{\hat{\theta}\hat{\phi}\hat{\theta}\hat{\phi}}. \quad (2.207)$$

The last equality follows from the symmetries of the Riemann tensor.

2.6 Gravity is a Lorentzian Connection on Spacetime

2.6.1 The Four Key Principles of General Relativity

As we discussed in the previous sections, the concept of spacetime in the form of a pseudo-Riemannian four-dimensional manifold is essential for Einstein's vision of gravity:

Premise I: Spacetime is a four-dimensional manifold M endowed with a global symmetric metric field g .

As we have seen, this guarantees that gravitational redshift is a natural consequence for this description of gravity.

In order to compare tangent spaces at neighboring events, we also need a **connection** on this manifold M . As in Riemannian geometry, this connection is required to be metric, so that the corresponding Christoffel symbols are uniquely given by derivatives of the metric elements. This is a basic postulate of Einstein's theory of gravity – one could construct more general theories of gravity which include torsion.

Physically speaking, we associate *observers* e_a ($a = 0, 1, 2, 3$), i.e. an orthonormal tetrad (or Vierbein field), satisfying³

$$g(e_a, e_b) = \eta_{ab}, \quad (2.208)$$

where η is the flat Minkowskian metric with signature $(+ - - -)$, or $(- + + +)$. An observer is a global orthonormal basis field in the tangent space of each event p , where e_0 is time-like and e_i ($i = 1, 2, 3$) are space-like. One could also construct null tetrads in order to define the geometry of the spacetime. The dual elements of e_a is a basis of the cotangent space T_p^* , denoted by Θ^a which define the metric $g = \eta_{ab} \Theta^a \otimes \Theta^b$. The definition of these observer fields is not unique, since any observer derived by means of a local Lorentz transformation Λ is also an observer

$$\bar{e}_a|_x = \Lambda_a^b(x) e_b|_x \quad , \quad \Lambda(x) \eta \Lambda^T(x) = \eta . \quad (2.209)$$

These are Lorentz transformations operating in the tangent space of each event.

For any pseudo-Riemannian manifold there is then a unique affine connection ∇ such that it is (i) torsion-free and (ii) metric. This particular connection is usually called the **Levi-Civita connection**, or pseudo-Riemannian connection.

Premise II: It is now one of the fundamental postulates of Einstein's theory of gravity that gravity is related to the Levi-Civita connection of the Lorentzian

³ In the following, the convention for indices is as follows: Greek indices are related to local coordinate systems, Latin indices a, b, c, \dots mark observer fields, Latin indices i, k, l, \dots specify spatial components.

manifold. This means in particular that there is no torsion associated with gravity!

The Einstein principle of equivalence states that effects of gravitation can be transformed away locally by suitably accelerated frames of reference (by going to local inertial coordinates). It also includes the stronger requirement

Premise III: Any physical interaction (other than gravitation) behaves in a local inertial frame as if gravitation were absent. Maxwell's equations, for example, will have their familiar forms as in SR.

This principle of equivalence allows us to extend any physical law that is expressed in a covariant way to curved spacetime. Ordinary derivatives are just replaced by covariant ones.

The energy-momentum tensor of a **perfect fluid** is then readily extended to a general spacetime just by replacing the flat metric η in terms of the metric field g (signature $(- +++)$)

$$T^{\mu\nu} = (\varrho + P)u^\mu u^\nu + Pg^{\mu\nu}. \quad (2.210)$$

As in special relativity, ϱ is the total energy density, P the pressure and u^μ the velocity field. Since in special relativity, the equations of motion follow from the divergence of this energy-momentum tensor, the above principle just requires the replacement of the ordinary derivatives by covariant ones

$$T_{;\nu}^{\mu\nu} = 0. \quad (2.211)$$

We already know of a symmetric $(0, 2)$ tensor, constructed from the Ricci tensor, which is automatically conserved: the **Einstein tensor**

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \quad (2.212)$$

which always obeys $\nabla_\mu G_\nu^\mu = 0$ (see last section). Einstein was therefore led to propose [141]

$$G_{\mu\nu} = \kappa T_{\mu\nu} \quad (2.213)$$

as field equations for the metric. This equation satisfies all of the obvious requirements; the right-hand side is a covariant expression of the energy and momentum density in the form of a symmetric and conserved $(0,2)$ tensor, while the left-hand side is a symmetric and conserved $(0,2)$ tensor constructed from the metric and its first and second derivatives. It only remains to see whether it actually reproduces gravity as we know it.

In order to prove this, we go to the Newtonian limit

$$g_{00} = 1 + h_{00}, \quad g_{0i} = 0, \quad g_{ik} = -\delta_{ik}. \quad (2.214)$$

As a source we take the energy-momentum tensor of a perfect fluid with vanishing pressure (sometimes called dust)

$$T_{\mu\nu} = \varrho u^\mu u^\nu. \quad (2.215)$$

In the rest frame of the fluid we have

$$u^\mu = (u^0, 0, 0, 0) \quad (2.216)$$

and therefore from the normalization $u^\mu u_\mu = 1$

$$u^0 = 1 + \frac{1}{2} h_{00} \simeq 1. \quad (2.217)$$

The trace of T is then

$$T = g^{00} T_{00} = T_{00} = \varrho. \quad (2.218)$$

On the other hand we find from the proposed field equations

$$R_{\mu\nu} = \kappa(T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu}), \quad (2.219)$$

since in four dimensions $R = -\kappa T$. This is essentially the same equation. The essential component predicts therefore

$$R_{00} = \frac{1}{2} \kappa \varrho. \quad (2.220)$$

We have to calculate $R_{00} = R^i_{0i0}$ with

$$R^i_{0j0} = \Gamma^i_{00,j} - \Gamma^i_{j0,0} + \Gamma^i_{j\alpha} \Gamma^{\alpha}_{00} - \Gamma^i_{0\alpha} \Gamma^{\alpha}_{j0}. \quad (2.221)$$

The second term is a time derivative which vanishes for static fields. The third and fourth terms are quadratic in the Christoffels, and we are only looking for first-order terms. So they can be neglected. From this we get

$$\begin{aligned} R_{00} &= R^i_{0i0} = \Gamma^i_{00,i} \\ &= \partial_i \left[\frac{1}{2} g^{i\alpha} (g_{\alpha 0,0} + g_{0\alpha,0} - g_{00,\alpha}) \right] \\ &= -\frac{1}{2} g^{ij} \partial_i \partial_j g_{00} = +\frac{1}{2} \delta^{ij} \partial_i \partial_j h_{00} = \frac{1}{2} \nabla^2 h_{00}. \end{aligned} \quad (2.222)$$

So we see that the time component of equation (2.219) predicts in the Newtonian limit the Poisson equation

$$\nabla^2 h_{00} = \kappa \varrho. \quad (2.223)$$

Since the gravitational redshift has already fixed the Newtonian limit of the metric element, $h_{00} = 2\Phi/c^2$, this is the correct Poisson equation provided the coupling constant is set to $\kappa = 8\pi G/c^4$. This defines the field equations of general relativity

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \quad (2.224)$$

G , of course, is the Newtonian gravitational constant. We have seen, Einstein's equations can also be written in the form for the Ricci tensor alone

$$R_{\mu\nu} = \frac{8\pi G}{c^4} (T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}). \quad (2.225)$$

They tell us in particular that **vacuum fields satisfy**

$$R_{\mu\nu} = 0. \quad (2.226)$$

Premise IV: In November 1915, Einstein postulated that the tensor $G_{\mu\nu}$ couples to the matter content of the Universe

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \quad (2.227)$$

where $T_{\mu\nu}$ is the symmetric energy-momentum tensor of all matter in the Universe (particles, baryons, galaxies, photons, neutrinos, quantum fields, as well as vacuum energy). As a consequence of the above properties, the divergence of the energy-momentum tensor vanishes identically

$$T^{\mu\nu}_{;\nu} = 0. \quad (2.228)$$

2.6.2 The Hilbert Action and Einstein's Field Equations

To increase your confidence that Einstein's equations as we have derived them are indeed the correct field equations for the metric, let's see how they can be derived from a more modern viewpoint, starting from an action principle. In fact the equations were first derived by Hilbert, not Einstein, and Hilbert did it using the action principle. But he had been inspired by Einstein's previous papers on the subject, and Einstein himself derived the equations independently, so they are rightly named after Einstein. The action, however, is rightly called the Hilbert action. The action should be the integral over spacetime of a Lagrange density (*Lagrangian* for short, although strictly speaking the Lagrangian is the integral over space of the Lagrange density)

$$S_H = \int \mathcal{L} d^4x. \quad (2.229)$$

The Lagrange density is a tensor density, which can be written as $\sqrt{-g}$ times a scalar. What scalars can we make out of the metric? Since we know that the metric can be set equal to its canonical form and its first derivatives set to zero at any one point, any nontrivial scalar must involve at least second derivatives of the metric. The Riemann tensor is of course made from second derivatives of the metric, and we argued earlier that the only independent scalar we could construct from the Riemann tensor was the Ricci scalar R . What we did not show, but is nevertheless true, is that any nontrivial tensor made from the metric and its first and second derivatives can be expressed in terms of the metric and the Riemann tensor. Therefore, the only independent scalar constructed from the metric, which is no higher than second order in its derivatives, is the Ricci scalar. Hilbert figured out that this was therefore the simplest possible choice for a Lagrangian, and proposed $\mathcal{L}_H = \sqrt{-g}R$, i.e. the **Hilbert action** reads as

$$S_H = \int \sqrt{-g} R d^4x. \quad (2.230)$$

The Hilbert action principle is valid in any dimension of the manifold. It can be formulated, for example, for 11D supergravity.

2.6.3 On the Cosmological Constant

Let us now consider a new matter action $L'_{\text{matter}} = L_{\text{matter}} - \Lambda/(8\pi G)$, where Λ is a real constant. The equation of motion for the matter does not change under this transformation, since Λ is constant. But the action now picks up an extra term proportional to Λ , which can be written in two different ways,

$$\begin{aligned} S &= -\frac{1}{16\pi G} \int R \sqrt{-g} d^4x + \int \left(L_{\text{matter}}(\Phi, \partial\Phi) - \frac{\Lambda}{8\pi G} \right) \sqrt{-g} d^4x \\ &= -\frac{1}{16\pi G} \int (R + 2\Lambda) \sqrt{-g} d^4x + \int L_{\text{matter}}(\Phi, \partial\Phi) \sqrt{-g} d^4x \end{aligned} \quad (2.231)$$

and Einstein's equations get modified. This simple manipulation has many backdrops in theoretical physics. It can be interpreted in different manners:

- The first interpretation is based on the first line of the above equations, it treats Λ as a **shift** in the matter Lagrangian, which in turn will lead to a shift in the matter Hamiltonian. This could be thought of as a **shift in the zero point energy** of the matter system. Such a constant shift in energy does not affect the dynamics of matter, while gravity picks up an extra contribution in the form of a new term $Q_{\mu\nu}$ in the energy-momentum tensor

$$R^\mu_\nu - \frac{1}{2} R \delta^\mu_\nu = 8\pi G(T^\mu_\nu + Q^\mu_\nu) \quad , \quad Q^\mu_\nu = \frac{\Lambda}{8\pi G} \delta^\mu_\nu. \quad (2.232)$$

- The second line in Eq (2.231) can be interpreted as a gravitational field, described by the Lagrangian of the form $L_{\text{grav}} \propto (1/G)(R+2\Lambda)$, interacting with matter. In

this interpretation, gravity is described by two constants, the Newton's constant G and the cosmological constant Λ . It is then natural to modify the left-hand side of Einstein's equations in the form of

$$R_v^\mu - \frac{1}{2} R \delta_v^\mu - \delta_v^\mu \Lambda = 8\pi G T_v^\mu. \quad (2.233)$$

In this interpretation, the spacetime is curved even in the absence of matter, $T_{\alpha\beta} = 0$, since the left-hand side does not admit flat spacetimes as solutions.

- It is even possible to consider a situation where both effects can occur. If gravitational theories are in fact described by the Lagrangian of the form $(R + 2\Lambda)$, then there is an intrinsic cosmological constant in nature, just as there is a Newtonian constant G in nature. If the matter Lagrangian contains energy densities which change due to the dynamics, then L_{matter} can pick up constant shifts during dynamical evolution. For this we consider a scalar field with the Lagrangian

$$L_\Phi = (1/2) \partial_\mu \Phi \partial^\mu \Phi - V(\Phi), \quad (2.234)$$

which has the energy-momentum tensor

$$T_v^\mu = \partial^\mu \Phi \partial_v \Phi - \delta_v^\mu \left(\frac{1}{2} \partial^\mu \Phi \partial_v \Phi - V(\Phi) \right). \quad (2.235)$$

For field configurations which are constant (e.g. at the minimum of the potential V), this contributes an energy-momentum tensor $T_v^\mu = V(\Phi_{\min}) \delta_v^\mu$, which has exactly the same form as a cosmological constant. It is then the combination of these two effects – of very different nature – which is relevant and the source will be

$$T_{\mu\nu}^{\text{eff}} = [V(\Phi_{\min}) + \Lambda/(8\pi G)] g_{\mu\nu}. \quad (2.236)$$

Φ_{\min} can change during the dynamical evolution, leading to a time-dependent cosmological constant.

The term $Q_{\mu\nu}$ in Einstein's equations behaves very peculiarly compared to the energy-momentum tensor of normal matter. $Q_v^\mu = \varrho_A \delta_v^\mu$ is in the form of an energy-momentum tensor of an ideal fluid with energy density ϱ_A and pressure $P_A = -\varrho_A$. Obviously, either the pressure or the energy density of this fluid must be negative.

Such an equation of state, $P = -\varrho$, also has another important implication in GR. The relative acceleration between two geodesics, \mathbf{g} , satisfies in GR the following equation (without proof)

$$\nabla \cdot \mathbf{g} = -4\pi G(\varrho + 3P). \quad (2.237)$$

The source of this relative acceleration between geodesics is $\varrho + 3P$ and not ϱ alone. This shows, as long as $\varrho + 3P > 0$, gravity remains attractive, while $\varrho + 3P < 0$ leads to repulsive forces. A positive cosmological constant therefore leads to repulsive gravity.

The condition $\varrho + 3P < 0$ has for a long time been considered to be unphysical. In fact, the so-called **strong energy condition** requires $\varrho + P \geq 0$ and $\varrho + 3P \geq 0$. This would mean that gravity is always attractive. In the meantime, we have however learnt that at least vacuum energy will violate the strong energy condition – and this is in fact realized in the present Universe: it seems that vacuum energy is the dominant contribution to the source of gravity in the present Universe.

2.6.4 Limits of General Relativity

General relativity is a classical field theory. In other field theories (electromagnetism, weak and strong interaction), there are well-understood procedures for **quantizing** these theories. For general relativity, however, these procedures run into technical and conceptual difficulties (infinities, nonrenormalizability, etc.).

The regime where observational effects of quantum gravity are expected to become important is far from the weak limit. In 1899, Planck noticed that his constant \hbar could be combined with Newton's constant G and the speed of light c to form a basic set of scales⁴

- **Planck mass** m_P , defined as

$$m_P = \sqrt{\frac{\hbar c}{G}} = 2.18 \times 10^{-5} \text{ g} = 1.22 \times 10^{19} \text{ GeV}/c^2, \quad (2.238)$$

- **Planck length** Λ_P , defined as

$$\Lambda_P = \frac{\hbar}{m_P c} = \sqrt{\frac{\hbar G}{c^3}} = 1.62 \times 10^{-33} \text{ cm}, \quad (2.239)$$

- **Planck time** t_P , defined as

$$t_P = \frac{\Lambda_P}{c} = \sqrt{\frac{\hbar G}{c^5}} = 5.34 \times 10^{-44} \text{ sec}. \quad (2.240)$$

Quantum gravity is expected to become important, when physical scales are of the order of the Planck scales, i.e. when masses $m \simeq m_P$, or when curvature radii $\simeq \Lambda_P$, or when time-scales in collapse situations and in the early Universe are $\simeq t_P$. A discussion of the Big Bang for time-scales shorter than $10 t_P$ is not meaningful. Quantum gravity is therefore important in the very early Universe and near the singularities of black holes, but certainly not in the Solar System.

At present, one of the deepest problems in theoretical physics is harmonizing the theory of general relativity, which describes gravitation and applies to large-scale structures (stars, planets, galaxies), with quantum mechanics, which describes the other three fundamental forces acting on the microscopic scale.

A fundamental lesson of general relativity is that there is no fixed spacetime background, as found in Newtonian mechanics and special relativity; the spacetime

⁴ $\hbar = 1.05 \times 10^{-27} \text{ g cm}^2 \text{ s}^{-1}$, $G = 6.67 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2}$, $c = 2.998 \times 10^{10} \text{ cm s}^{-1}$.

geometry is dynamical. While easy to grasp in principle, this is the hardest idea to understand about general relativity, and its consequences are profound and not fully explored, even at the classical level. To a certain extent, general relativity can be seen to be a relational theory, in which the only physically relevant information is the relationship between different events in spacetime.

On the other hand, quantum mechanics has depended since its invention on a fixed background (nondynamical) structure. In the case of quantum mechanics, it is time that is given and not dynamical, just as in Newtonian classical mechanics. In relativistic quantum field theory, just as in classical field theory, Minkowski spacetime is the fixed background of the theory. Finally, string theory started out as a generalization of quantum field theory where instead of point particles, string-like objects propagate in a fixed spacetime background; condensation of particular vibration modes of strings is equivalent to a modification of the original background.

In the 1920s the French mathematician Elie Cartan formulated Einstein's theory in the language of bundles and connections, a generalization of Riemann's geometry to which Cartan made important contributions. The so-called Einstein–Cartan theory of gravity not only reformulated, but also generalized general relativity, and allowed spacetimes with torsion as well as curvature. In Cartan's geometry of bundles, the concept of parallel transport is more fundamental than that of distance, the centerpiece of Riemannian geometry. A similar conceptual shift occurs between the invariant interval of Einstein's general relativity and the parallel transport of Einstein–Cartan theory.

In the 1960s, Roger Penrose explored the idea of space arising from a quantum combinatorial structure. His investigations resulted in the development of spin networks. Because this was a quantum theory of the rotational group and not the Lorentz group, Penrose went on to develop twistors.

In 1986, Abhay Ashtekar [44,45] reformulated Einstein's field equations of general relativity using what have come to be known as Ashtekar variables, a particular flavor of Einstein–Cartan theory with a complex connection. Using this reformulation, he was able to quantize gravity using well-known techniques from quantum gauge field theory. In the Ashtekar formulation, the fundamental objects are a rule for parallel transport (technically, a connection) and a coordinate frame (called a vierbein) at each point.

Loop quantum gravity [12] is the best one can do so far in trying to understand quantum spacetime, from a nonperturbative, background-independent point of view. Theoretically, we have reasons to suspect that this approach could represent a consistent quantum theory with the correct classical limit, but there are also some worrying contrary indications. The theory yields a definite physical picture of quantum spacetime and definite quantitative predictions, but a systematic way of extracting physical information is still lacking. Experimentally, there is no support to the theory, neither direct nor indirect. The spectra of area and volume computed in the theory could or could not be physically correct.

The most remarkable physical result obtained from loop quantum gravity is evidence for a **physical (quantum) discreteness of space at the Planck scale**.

This is manifested in the fact that certain operators corresponding to the measurement of geometrical quantities, in particular area and volume, have discrete spectrum. According to the standard interpretation of quantum mechanics, this means that the theory predicts that a physical measurement of an area or a volume will necessarily yield quantized results. Since the smallest eigenvalues are of Planck scale, this implies that there is no way of observing areas or volumes smaller than Planck scale. Space comes in “quanta” in the same manner as the energy of an oscillator. The spectra of the area and volume operators have been computed with much detail in loop quantum gravity. These spectra have a complicated structure, and they constitute detailed quantitative physical predictions of loop quantum gravity on Planck scale physics. If we had experimental access to Planck scale physics, they would allow the theory to be empirically tested in great detail.

2.7 Gravitational Waves

In the previous sections we have formulated the Einstein field equations in a neat and compact form. This should not hide the fact that the Einstein tensor G is in fact a complicated function of the metric g and its first and second derivatives. Due to the symmetry of the Einstein tensor and the energy momentum tensor the field equations represent 10 coupled, nonlinear partial differential equations.

It therefore came as quite a surprise when Karl Schwarzschild found a nontrivial, analytic solution to these equations just some months after their publication. Since then many analytic solutions have been found and a whole branch of the studies of general relativity is concerned with their classification. Enormous insight into the structure of general relativity has been gained from these analytic solutions, but due to the complexity of the field equations these solutions are normally idealized and restricted by symmetry assumptions. In order to obtain accurate descriptions of astrophysically relevant scenarios one may therefore have to go beyond purely analytic studies.

2.7.1 The Geodesic Deviation – Relativistic Tidal Forces

Let us consider a congruence of time-like geodesics $x^\mu(\tau, s)$ given by their proper time τ and a coordinate s perpendicular to the geodesics. With this we can define the four-velocity $U^\mu = \partial x^\mu / \partial \tau$ and a space-like connecting vector $X^\mu = \partial x^\mu / \partial s$. These two fields obviously commute

$$[U, X] = 0. \quad (2.241)$$

Therefore

$$[U, X] = \nabla_U X - \nabla_X U = 0. \quad (2.242)$$

Let us now calculate the transverse acceleration

$$\begin{aligned}
 \frac{D^2 X^\mu}{d\tau^2} &= U^\varrho \nabla_\varrho (U^\sigma \nabla_\sigma X^\mu) \\
 &= U^\varrho \nabla_\varrho (X^\sigma \nabla_\sigma U^\mu) \\
 &= (U^\varrho \nabla_\varrho X^\sigma) (\nabla_\sigma U^\mu) + U^\varrho X^\sigma \nabla_\varrho \nabla_\sigma U^\mu \\
 &= (X^\sigma \nabla_\sigma U^\varrho) (\nabla_\varrho U^\mu) + U^\varrho X^\sigma \left[\nabla_\sigma \nabla_\varrho U^\mu + R^\mu_{\nu\varrho\sigma} U^\nu \right] \\
 &= (X^\sigma \nabla_\sigma U^\varrho) (\nabla_\varrho U^\mu) + X^\sigma \nabla_\sigma (U^\varrho \nabla_\varrho U^\mu) \\
 &\quad - (X^\sigma \nabla_\sigma U^\varrho) (\nabla_\varrho U^\mu) + R^\mu_{\nu\varrho\sigma} U^\nu U^\varrho X^\sigma \\
 &= R^\mu_{\nu\varrho\sigma} U^\nu U^\varrho X^\sigma. \tag{2.243}
 \end{aligned}$$

The two first terms cancel each other and the second term vanishes for geodetic motion. The transverse acceleration is the generalization of the tidal force equation in Newtonian mechanics.

2.7.2 Gravity Wave Experiments

A particularly important area of research connected with general relativity that has emerged in recent years concerns the detection of gravitational waves. In analogy to the prediction of electromagnetic waves by the Maxwell equations of electrodynamics, the Einstein field equations admit radiative solutions with a characteristic propagation speed given by the speed of light. In contrast to electromagnetic waves, which will only exert a force on charged particles, gravitational waves will have an effect on any kind of matter. This effect is best illustrated by considering a set of freely moving test particles arranged in a circle (see Fig. 2.4). If a gravitational wave falls onto the plane spanned by these test particles (i.e. coming out of the picture), it will cause the particles to move. Gravitational waves are known to possess two states of polarization called plus (+) and cross (\times) and the induced motion of the particles will depend on the polarization.

Due to the weak coupling constant of the gravitational interaction, the particle motion induced by gravitational waves will have an extremely small amplitude in reality. If one considers for example a bar of a length of several kilometers, estimates have shown that the detection of gravitational waves requires one to measure changes in length orders of magnitude smaller than the diameter of an atomic nucleus. Even though attempts to detect gravitational radiation go back to the work of Joe Weber in the early sixties, it is only the recent advance of computer and laser technology that provides scientists with a realistic chance of success. The current generation of gravitational wave detectors GEO-600, LIGO, TAMA and VIRGO that have been constructed for this purpose are complex multinational collaborations and have recently gone online or are expected to go online in the near future. The following image gives an overview over all wave experiments presently in operation or under construction (Fig. 2.3).

These detectors essentially consist of two bars with a length of hundreds of meters or kilometers arranged in the shape of an “L”. Variations in length of these

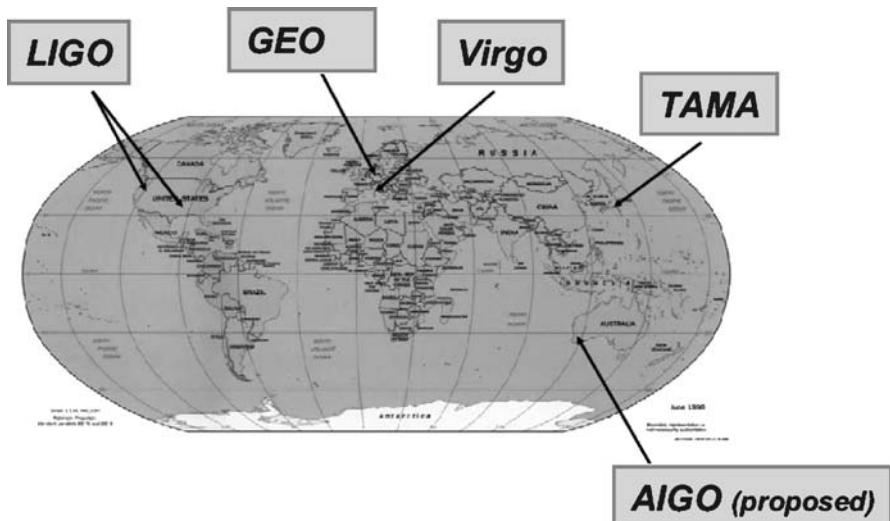


Fig. 2.3. Gravity wave experiments based on laser interferometric techniques

bars are measured by lasers to very high accuracy. Due to the extreme smallness of the signals, however, the accumulation of data over several years will be necessary to improve the chances of a positive identification of gravitational wave signals from extragalactic sources.

Confidence in the existence of gravitational waves has been significantly boosted by the Nobel prize winning discovery of the binary neutron star system PSR1913+16 (sometimes referred to as the *Hulse and Taylor pulsar*). The spin-down of this system has been found to agree remarkably well with the energy loss predicted by general relativity due to the emission of gravitational waves and is generally accepted as indirect proof of the existence of gravitational radiation (will be discussed in Sect. 6.5).

In order to simplify the enormous task of detecting gravitational waves, it is vital to obtain information about the structure of the signals one is looking for. It is necessary for this purpose to accurately model the astrophysical scenarios that are considered likely sources of gravitational waves and extract the corresponding signals from these models. According to Birkhoff's theorem the Schwarzschild solution, which describes a static, spherically symmetric vacuum spacetime, is the only spherically symmetric, asymptotically flat solution to the Einstein vacuum field equations. As a consequence a spherically symmetric spacetime, even if it contains a radially pulsating object, will necessarily have an exterior static region and be nonradiating. It is necessary, therefore, to use less restrictive symmetry assumptions in the modelling of astrophysical sources of gravitational waves. In fact the most promising sources of gravitational waves currently under consideration are the in-spiralling and merger of two compact bodies (neutron stars or black holes) and complicated oscillation modes of neutron stars that increase in amplitude due to

the emission of gravitational waves by extracting energy from the rotation of the star. Even though a great deal of information about these scenarios has been gained from approximative studies, such as the post-Newtonian formalism or the use of perturbative techniques, a detailed simulation will require the solution of the Einstein equations in three dimensions. The complicated structure of the corresponding models in combination with the enormous advance in computer technology has given rise to numerical relativity, the computer based generation of solutions to Einstein's field equations.

2.7.3 The Nature of Gravitational Waves

Gravitational waves are ripples in the fabric of space and time produced by violent events in the distant Universe, for example by the collision of two black holes or by the cores of supernova explosions. Gravitational waves are emitted by accelerating masses much as electromagnetic waves are produced by accelerating charges. These ripples in the spacetime fabric travel to Earth, bringing with them information about their violent origins and about the nature of gravity.

Linearized Gravity

The weakness of the gravitational field is once again expressed as our ability to decompose the metric into the flat Minkowski metric plus a small perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad , \quad |h_{\mu\nu}| \ll 1. \quad (2.244)$$

Under this condition, the inverse metric is simply given by

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}, \quad (2.245)$$

where $h^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}$. We can raise and lower indices just by using the flat Minkowski metric η . For this reason we may consider $h_{\mu\nu}$ as a symmetric tensor of second rank defined on Minkowski space.

We want to find the equation of motion obeyed by the perturbations h , which come by examining Einstein's equations to first order. We begin with the Christoffel symbols, which are given by

$$\Gamma_{\mu\nu}^\rho = \frac{1}{2}\eta^{\rho\sigma}(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}). \quad (2.246)$$

Since the connection coefficients are first-order quantities, the only contribution to the Riemann tensor will come from the derivatives of the Γ 's, not the Γ^2 terms. Lowering an index for convenience, we obtain

$$\begin{aligned} R_{\mu\nu\theta\sigma} &= \eta_{\mu\lambda}\partial_\theta\Gamma_{\nu\sigma}^\lambda - \eta_{\mu\lambda}\partial_\sigma\Gamma_{\nu\theta}^\lambda \\ &= \frac{1}{2}[\partial_\theta\partial_\nu h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\theta} - \partial_\sigma\partial_\nu h_{\mu\theta} - \partial_\theta\partial_\mu h_{\nu\sigma}]. \end{aligned} \quad (2.247)$$

The Ricci tensor is obtained by contracting over μ and ϱ , giving

$$R_{\mu\nu} = \frac{1}{2}[\partial_\sigma \partial_\nu h^\sigma_\mu + \partial_\sigma \partial_\mu h^\sigma_\nu - \partial_\mu \partial_\nu h - \square h_{\mu\nu}]. \quad (2.248)$$

Here we have defined the trace of the perturbation, $h = \eta^{\mu\nu} h_{\mu\nu}$, and the d'Alembertian operator in Minkowski space $\square = \eta^{\mu\nu} \partial_\mu \partial_\nu = \partial_t^2 - \partial_x^2 - \partial_y^2 - \partial_z^2$. Finally, we obtain the Ricci scalar

$$R = \partial_\mu \partial_\nu h^{\mu\nu} - \square h. \quad (2.249)$$

Putting all this together, we obtain the Einstein tensor

$$\begin{aligned} G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2}R\eta_{\mu\nu} \\ &= \frac{1}{2}[\partial_\sigma \partial_\nu h^\sigma_\mu + \partial_\sigma \partial_\mu h^\sigma_\nu - \partial_\mu \partial_\nu h \\ &\quad - \square h_{\mu\nu} - \eta_{\mu\nu} \partial_\varrho \partial_\lambda h^{\varrho\lambda} + \eta_{\mu\nu} \square h]. \end{aligned} \quad (2.250)$$

The linearized field equations are then

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \quad (2.251)$$

where $T_{\mu\nu}$ is the energy-momentum tensor calculated in zeroth order from h . We do not include higher-order corrections to the energy-momentum tensor, because the amount of energy and momentum must itself be small for the weak-field limit to apply. In other words, the lowest nonvanishing order in T is automatically of the same order of magnitude as the perturbation. Notice that the conservation law to lowest order is simply $\partial_\mu T^{\mu\nu} = 0$. We will most often be concerned with the vacuum equations, which as usual are just $R_{\mu\nu} = 0$, where R is given by (2.248).

On Gauge Invariance

With the linearized field equations in hand, we are almost prepared to set about solving them. First, however, we should deal with the important issue of **gauge invariance**. This issue arises because the demand that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ does not completely specify the coordinate system on spacetime; there may be other coordinate systems, in which the metric can still be written as the Minkowski metric plus a small perturbation, but the perturbation will be different. Thus, the **decomposition of the metric into a flat background plus a perturbation is not unique**.

The notion that the linearized theory can be thought of as one governing the behavior of tensor fields on a flat background can be formalized in terms of a “background spacetime” M_0 , a “physical spacetime” M_p , and a diffeomorphism $\Phi : M_0 \rightarrow M_p$. As manifolds M_0 and M_p are “the same” (since they are diffeomorphic), but we imagine that they possess some different tensor fields; on M_0 we have defined the flat Minkowski metric η , while on M_p we have some metric g which

obeys Einstein's equations. (We imagine that M_0 is equipped with coordinates x and M_p is equipped with coordinates y , although these will not play a prominent role.) The diffeomorphism Φ allows us to move tensors back and forth between the background and physical spacetimes. Since we would like to construct our linearized theory as one taking place on the flat background spacetime, we are interested in the pullback ($\Phi^* g$) of the physical metric. We can define the perturbation as the difference between the pulled-back physical metric and the flat one:

$$h_{\mu\nu} = (\Phi^* g)_{\mu\nu} - \eta_{\mu\nu}. \quad (2.252)$$

From this definition, there is no reason for the components of h to be small; however, if the gravitational fields on M_p are weak, then for some diffeomorphisms Φ we will have $|h_{\mu\nu}| \ll 1$. We therefore limit our attention only to those diffeomorphisms for which this is true. Then the fact that g obeys Einstein's equations on the physical spacetime means that h will obey the linearized equations on the background spacetime (since Φ , as a diffeomorphism, can be used to pull back Einstein's equations themselves).

In this language, the issue of gauge invariance is simply the fact that there are a large number of permissible diffeomorphisms between M_0 and M_p (where “permissible” means that the perturbation is small). Consider a vector field $\xi^\mu(x)$ on the background spacetime. This vector field generates a one-parameter family of diffeomorphisms $\Psi_\epsilon : M_0 \rightarrow M_0$. For ϵ sufficiently small, if Φ is a diffeomorphism for which the perturbation defined by h is small than so will $(\Phi \cdot \Psi_\epsilon)$ be, although the perturbation will have a different value. Specifically, we can define a family of perturbations parameterized by ϵ

$$h_{\mu\nu}^{(\epsilon)} = [(\Phi \cdot \Psi_\epsilon)^* g]_{\mu\nu} - \eta_{\mu\nu} = [\Psi_\epsilon(\Phi^* g)]_{\mu\nu} - \eta_{\mu\nu}. \quad (2.253)$$

The second equality is based on the fact that the pullback under a composition is given by the composition of the pullbacks in the opposite order, which follows from the fact that the pullback itself moves things in the opposite direction from the original map. Plugging in the relation for h , we find

$$h_{\mu\nu}^{(\epsilon)} = \Psi_\epsilon^*(h + \eta)_{\mu\nu} - \eta_{\mu\nu} = \Psi_\epsilon^*(h)_{\mu\nu} + \Psi_\epsilon^*(\eta)_{\mu\nu} - \eta_{\mu\nu}, \quad (2.254)$$

since the pullback of the sum of two tensors is the sum of the pullbacks. Now we use our assumption that ϵ is small; in this case $\Psi_\epsilon^*(h)$ will be equal to h to lowest order, while the other two terms give us a Lie derivative

$$\begin{aligned} h_{\mu\nu}^{(\epsilon)} &= \Psi_\epsilon^*(h) + \epsilon \left[\frac{\Psi_\epsilon^*(\eta_{\mu\nu}) - \eta_{\mu\nu}}{\epsilon} \right] \\ &= h_{\mu\nu} + \epsilon L_\xi \eta_{\mu\nu}. \end{aligned} \quad (2.255)$$

Since the background metric is flat, we therefore find

$$h_{\mu\nu}^{(\epsilon)} = h_{\mu\nu} + \epsilon (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu).$$

(2.256)

This formula represents the change of the metric perturbation under an infinitesimal diffeomorphism along the vector field $\epsilon\xi^\mu$: **this is called a gauge transformation in linearized theory.**

The infinitesimal diffeomorphisms Ψ_ϵ provide a different representation of the same physical situation, while maintaining our requirement that the perturbation be small. Therefore, the above result tells us what kind of metric perturbations denote physically equivalent spacetimes – those related to each other by $\epsilon(\partial_\mu\xi_\nu + \partial_\nu\xi_\mu)$, for some vector field ξ^μ . The invariance of our theory under such transformations is analogous to traditional gauge invariance of electromagnetism under $A_\mu \rightarrow A_\mu + \partial_\mu\lambda$. (The analogy is different from the previous analogy we drew with electromagnetism, relating local Lorenz transformations in the orthonormal-frame formalism to changes of basis in an internal vector bundle.) In electromagnetism the invariance comes about because the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is left unchanged by gauge transformations; similarly, we find that the transformation (2.256) changes the linearized Riemann tensor by

$$\delta R_{\mu\nu\rho\sigma} = 0. \quad (2.257)$$

Our abstract derivation of the appropriate gauge transformation for the metric perturbation is verified by the fact that it **leaves the curvature (and hence the physical spacetime) unchanged.**

Gauge invariance can also be understood from the slightly more lowbrow, but considerably more direct route of infinitesimal coordinate transformations. Our diffeomorphism Ψ_ϵ can be thought of as changing coordinates from x^μ to $x^\mu - \epsilon\xi^\mu$. (The minus sign, which is unconventional, comes from the fact that the “new” metric is pulled back from a small distance forward along the integral curves, which is equivalent to replacing the coordinates by those a small distance backward along the curves.) Following through the usual rules for transforming tensors under coordinate transformations, you can derive precisely (2.256) – although you have to cheat somewhat by equating components of tensors in two different coordinate systems.

2.7.4 Degrees of Freedom

The metric perturbation $h_{\mu\nu}$ is a symmetric (0,2) tensor on Minkowski spacetime. This means, under spatial rotations the 00 component is a scalar, the 0*i* component form a three-vector, and the *ij* components form a two-index symmetric spatial tensor. Each spatial tensor can be decomposed into a trace and a trace-free part (in group representations this corresponds to irreducible representations of the rotational group $SO(3)$). We therefore write $h_{\mu\nu}$ as

$$h_{00} = 2\Phi \quad (2.258)$$

$$h_{0i} = -w_i \quad (2.259)$$

$$h_{ij} = 2\Psi\delta_{ij} - 2s_{ij}. \quad (2.260)$$

Ψ denotes the trace of h_{ij} , and s_{ij} is traceless

$$\Psi = \frac{1}{6} \delta_{ij} h_{ij} \quad (2.261)$$

$$s_{ij} = -\frac{1}{2} \left(h_{ij} - \frac{1}{3} \delta^{kl} h_{kl} \delta_{ij} \right). \quad (2.262)$$

The entire metric can thus be written as⁵

$$ds^2 = (1 + 2\Phi) dt^2 - w_i(dt dx^i + dx^i dt) - [(1 - 2\Psi)\delta_{ij} - 2s_{ij}] dx^i dx^j. \quad (2.263)$$

Here we have not yet chosen a gauge, we just have conveniently decomposed the metric perturbations into two scalar modes, one vector mode and a tensor mode, adding to 10 independent components of the perturbation $h_{\mu\nu}$.

To get a feeling for the physical interpretation of these modes, we consider the motion of test particles as described by the geodesic equation. For this we need the Christoffel symbols

$$\Gamma_{00}^0 = \partial_0 \Phi \quad (2.264)$$

$$\Gamma_{00}^i = \partial_i \Phi + \partial_0 w_i \quad (2.265)$$

$$\Gamma_{j0}^0 = \partial_j \Phi \quad (2.266)$$

$$\Gamma_{j0}^i = \frac{1}{2} [\partial_j w_i - \partial_i w_j + \partial_0 h_{ij}] \quad (2.267)$$

$$\Gamma_{jk}^0 = -\frac{1}{2} [\partial_j w_k + \partial_k w_j - \partial_0 h_{ij}] \quad (2.268)$$

$$\Gamma_{jk}^i = \frac{1}{2} [\partial_j h_{ki} + \partial_k h_{ji} - \partial_i h_{jk}]. \quad (2.269)$$

Here we use $h_{ij} = -2s_{ij} + 2\Psi\delta_{ij}$. We decompose the four-momentum $p^\mu = dx^\mu/d\lambda$, where $\lambda = \tau/m$ for massive particles, in terms of the energy E and the three-velocity $v^i = dx^i/dt$

$$p^0 = \frac{dt}{d\lambda} = E \quad , \quad p^i = Ev^i. \quad (2.270)$$

Then we write the geodesic equation as though a force would act on the particles

$$\frac{dp^\mu}{dt} = -\Gamma_{\varrho\sigma}^\mu \frac{p^\varrho p^\sigma}{E}. \quad (2.271)$$

For $\mu = 0$ we get the energy evolution

$$\frac{dE}{dt} = -E \left[\partial_0 \Phi + 2(\partial_k \Phi)v^k - \frac{1}{2} (\partial_j w_k + \partial_k w_j - \partial_0 h_{jk}) v^j v^k \right]. \quad (2.272)$$

⁵ This ansatz can easily be generalized to cosmological spacetimes in order to describe general perturbations evolving under the expansion of the Universe.

The spatial components $\mu = i$ become

$$\begin{aligned} \frac{1}{E} \frac{dp^i}{dt} = & -\partial_i \Phi - \partial_0 w_i - (\partial_i w_j - \partial_j w_i + 2\partial_0 h_{ij}) v^j \\ & - \frac{1}{2} (\partial_j h_{ki} + \partial_k h_{ji} - \partial_i h_{jk}) v^j v^k. \end{aligned} \quad (2.273)$$

For a physical interpretation we introduce the **gravitoelectric and gravitomagnetic** field in terms of scalar and vector potentials, where the three-vector \mathbf{w} acts as a vector potential

$$G^i = -\partial_i \Phi - \partial_0 w_i \quad (2.274)$$

$$H^i = (\nabla \times \mathbf{w})^i = \epsilon^{ijk} \partial_j w_k. \quad (2.275)$$

Then we can write

$$\frac{1}{E} \frac{dp^i}{dt} = G^i + (\mathbf{v} \times \mathbf{H})^i - 2(\partial_0 h_{ij}) v^j - \frac{1}{2} (\partial_j h_{ki} + \partial_k h_{ji} - \partial_i h_{jk}) v^j v^k.$$

$$(2.276)$$

The first two terms on the right-hand side describe how the test particle responds to the scalar and vector perturbations Φ and w_i in a way reminiscent of the Lorenz force in electromagnetism. We also find couplings to the spatial perturbations h_{ij} of linear and quadratic order in the velocity.

Einstein's Equations

We can now decompose the Riemann tensor in our variables⁶

$$R_{0j0l} = \partial_j \partial_l \Phi + \partial_0 \partial_{(j} w_{l)} - \frac{1}{2} \partial_0^2 h_{jl} \quad (2.279)$$

$$R_{0jkl} = \partial_j \partial_{[k} w_{l]} - \partial_0 \partial_{[k} h_{l]k} \quad (2.280)$$

$$R_{ijkl} = \partial_j \partial_{[k} h_{l]i} - \partial_i \partial_{[k} h_{l]j}. \quad (2.281)$$

To obtain the Ricci tensor we contract with the flat metric η

$$R_{00} = \nabla^2 \Phi + \partial_0 \partial_k w^k + 3\partial_0^2 \Psi \quad (2.282)$$

$$R_{0j} = -\frac{1}{2} \nabla^2 w_j + \frac{1}{2} \partial_j \partial_k w^k + 2\partial_0 \partial_j \Psi + \partial_0 \partial_k s^k_j \quad (2.283)$$

$$R_{ij} = -\partial_i \partial_j (\Phi - \Psi) - \partial_0 \partial_{(i} w_{j)} + \square \Psi \delta_{ij} - \square s_{ij} + 2\partial_k \partial_{(i} s^k_{j)}, \quad (2.284)$$

⁶ We use the notation

$$2\partial_{(j} w_{k)} = \partial_j w_k + \partial_k w_j \quad (2.277)$$

$$2\partial_{[j} w_{k]} = \partial_j w_k - \partial_k w_j. \quad (2.278)$$

where $\nabla^2 = \delta^{ij} \partial_i \partial_j$ is the flat Laplacian. Finally, we can calculate the Einstein tensor

$$G_{00} = 2\nabla^2\Psi + \partial_k \partial_l s^{kl} \quad (2.285)$$

$$G_{0j} = -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j \partial_k w^k + 2\partial_0 \partial_j \Psi + \partial_0 \partial_k s_j^k \quad (2.286)$$

$$\begin{aligned} G_{ij} &= (\delta_{ij} \nabla^2 - \partial_i \partial_j)(\Phi - \Psi) + \delta_{ij} \partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} \\ &\quad + 2\delta_{ij} \partial_0^2 \Psi - \square s_{ij} + 2\partial_k \partial_{(i} s_{j)}^k - \delta_{ij} \partial_k \partial_m s^{km}. \end{aligned} \quad (2.287)$$

With this decomposition we see that in fact four equations are just constraint equations and do not present true dynamical evolution equations. To see this we start with the first equation which can be written as

$$\nabla^2 \Psi = 4\pi G T_{00} - \frac{1}{2} \partial_k \partial_m s^{km}. \quad (2.288)$$

This is an equation for Ψ with no time derivatives. If we know what are T_{00} and s_{ij} are doing all the time, the potential Ψ is uniquely determined by boundary conditions. Ψ is therefore not a propagating degree of freedom, it will be determined by the energy-momentum tensor and the strain. Next we consider the $0i$ equation, which we write as

$$(\delta_{jk} \nabla^2 - \partial_j \partial_k) w^k = -16\pi G T_{0j} + 4\partial_0 \partial_j \Psi + 2\partial_0 \partial_k s_j^k. \quad (2.289)$$

This is an equation for the vector field w^j which also does not contain time derivatives. Finally, the ij equation is

$$\begin{aligned} (\delta_{ij} \nabla^2 - \partial_i \partial_j) \Phi &= 8\pi G T_{ij} + (\delta_{ij} \nabla^2 - \partial_i \partial_j - 2\delta_{ij} \partial_0^2) \Psi \\ &\quad - \delta_{ij} \partial_0 \partial_k w^k + \partial_0 \partial_{(i} w_{j)} + \square s_{ij} - 2\partial_k \partial_{(i} s_{j)}^k - \delta_{ij} \partial_k \partial_m s^{jm}. \end{aligned} \quad (2.290)$$

Once again, there are not time derivatives acting on Φ , which is therefore determined from the other fields.

The only propagating degrees of freedom in Einstein's theory are those in the strain tensor s_{ij} . In terms of fields, which depend on the behavior under spatial rotations we may classify the scalars Ψ and Φ as spin-0, the vector w_i as spin-1 and the strain tensor as spin-2 degrees of freedom. Only the spin-2 degree of freedom is a true dynamical mode in general relativity.

Transverse Gauge

The different metric components of $h_{\mu\nu}$ will transform under a general gauge transformation generated by a vector field ξ^μ as

$$\Phi \rightarrow \Phi + \partial_0 \xi^0 \quad (2.291)$$

$$w_i \rightarrow w_i + \partial_0 \xi^i - \partial_i \xi^0 \quad (2.292)$$

$$\Psi \rightarrow \Psi - \frac{1}{3} \partial_i \xi^i \quad (2.293)$$

$$s_{ij} \rightarrow s_{ij} - \partial_{(i} \xi_{j)} - \frac{1}{3} \partial_k \xi^k \delta_{ij}. \quad (2.294)$$

First we consider the transverse gauge. This is closely related to the Coulomb gauge of electromagnetism, $\partial_i A^i = 0$. Similarly, we fix the strain by means of

$$\partial_i s^{ij} = 0, \quad (2.295)$$

by choosing ξ^j to satisfy

$$\nabla^2 \xi^j + \frac{1}{3} \partial_j \partial_i \xi^i = -2 \partial_i s^{ij}. \quad (2.296)$$

The value of ξ^0 is still undetermined. We can choose this term for the condition $\partial_i w^i = 0$ by means of

$$\nabla^2 \xi^0 = \partial_i w^i + \partial_0 \partial_i \xi^i. \quad (2.297)$$

With this gauge, Einstein's equations become

$$G_{00} = 2\nabla^2 \Psi = 8\pi G T_{00} \quad (2.298)$$

$$G_{0j} = -\frac{1}{2} \nabla^2 w_j + 2\partial_0 \partial_j \Psi = 8\pi G T_{0j} \quad (2.299)$$

$$G_{ij} = (\delta_{ij} \nabla^2 - \partial_i \partial_j)(\Phi - \Psi) \\ - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij} \partial_0^2 \Psi - \square s_{ij} = 8\pi G T_{ij}. \quad (2.300)$$

2.7.5 Gravitational Wave Solutions

Let us consider now the **transverse gauge**, by neglecting source terms, $T_{\mu\nu} = 0$. Then the 00 equation is

$$\nabla^2 \Psi = 0. \quad (2.301)$$

For suitable boundary conditions we can achieve $\Psi = 0$ everywhere. The $0i$ component is then

$$\nabla^2 w_i = 0, \quad (2.302)$$

which again implies $w_i = 0$. We turn next to the trace of the ij component with the above values

$$\nabla^2 \Phi = 0, \quad (2.303)$$

which also implies $\Phi = 0$.

We are then left with the trace-free part of the ij equation

$\square s_{ij} = 0,$

(2.304)

which becomes a wave equation for the traceless strain tensor. It is convenient to express the metric tensor in this **transverse traceless gauge**

$$h_{\mu\nu}^{TT} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -2s_{ij} & 0 \\ 0 & -2s_{ij} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.305)$$

This quantity is purely spatial, traceless and transverse, i.e.

$$h_{00}^{TT} = 0 \quad (2.306)$$

$$\eta^{\mu\nu} h_{\mu\nu}^{TT} = 0 \quad (2.307)$$

$$\partial^\mu h_{\mu\nu}^{TT} = 0. \quad (2.308)$$

In analogy to electromagnetism, plane waves are solutions of this equation

$$h_{\mu\nu}^{TT} = A_{\mu\nu} \exp(ik \cdot x), \quad (2.309)$$

where $A_{\mu\nu}$ is a constant symmetric $(0, 2)$ tensor, which is purely spatial and traceless

$$A_{0\mu} = 0, \quad \eta^{\mu\nu} A_{\mu\nu} = 0. \quad (2.310)$$

The constant k -vector is the wavevector with $k_\mu k^\mu = 0$. The plane wave (2.309) is a solution of the linearized equation, provided the wavevector is null. This means that **gravitational waves propagate with the speed of light**. Any superposition of plane waves is also a solution.

The condition of transversality means that

$$\partial^\mu h_{\mu\nu}^{TT} = ik^\mu A_{\mu\nu} \exp(ik \cdot x) = 0 \quad (2.311)$$

or that

$$k^\mu A_{\mu\nu} = 0. \quad (2.312)$$

We now consider a wave travelling in the z -direction, i.e.

$$k^\mu = (\omega, 0, 0, k^3) = (\omega, 0, 0, \omega). \quad (2.313)$$

In this case, the transversality requires that $A_{3\nu} = 0$. The only nonzero components are therefore $A_{11}, A_{12}, A_{21}, A_{22}$. But $A_{\mu\nu}$ is traceless and symmetric, i.e. of the form

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & -A_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (2.314)$$

For a plane wave travelling in the z -direction, the two amplitudes A_{11} and A_{12} completely characterize the wave.

For getting a feeling what happens if a wave passes by, we consider the motion of test particles in the presence of the gravitational field represented by the wave. For this we consider the relative motion of nearby particles with four-velocities described

by the vector field U^μ . Nearby geodesics are then given in terms of a separation vector X^μ , which satisfies the equation of geodesic deviation

$$\frac{D^2}{d\tau^2} X^\mu = R_{\nu\varrho\sigma}^\mu U^\nu U^\varrho X^\sigma. \quad (2.315)$$

The four-velocity is simply given by

$$U^\mu = (1, 0, 0, 0). \quad (2.316)$$

Therefore, we only need to compute the Riemann tensor $R_{00\sigma}^\mu$ which is given by

$$R_{\mu 00\sigma} = \frac{1}{2} (\partial_0^2 h_{\mu\sigma}^{TT} + \partial_\sigma \partial_\mu h_{00}^{TT} - \partial_\sigma \partial_0 h_{\mu 0}^{TT} - \partial_\mu \partial_0 h_{\sigma 0}^{TT}). \quad (2.317)$$

But with $h_{\mu 0}^{TT} = 0$, we get simply

$$R_{\mu 00\sigma} = \frac{1}{2} \partial_0^2 h_{\mu\sigma}^{TT}. \quad (2.318)$$

For slowly moving particles we have in lowest order $\tau = t = x^0$, so the equation of geodesic deviation becomes

$$\frac{\partial^2}{\partial t^2} X^\mu = \frac{1}{2} \partial_0^2 h_{\sigma}^{TT,\mu} X^\sigma. \quad (2.319)$$

For our plane wave this means that only X^1 and X^2 will be affected – **the test particles are only disturbed in directions perpendicular to the wavevector**. This is similar to electromagnetism, where the electric and magnetic fields are orthogonal in a plane wave.

Our plane wave is characterized by two amplitudes which are denoted for convenience as follows

$$h_+ = A_{11} \quad (2.320)$$

$$h_\times = A_{12}, \quad (2.321)$$

so that the amplitude tensor has the form

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & h_+ & h_\times \\ 0 & h_\times & -h_+ \\ 0 & 0 & 0 \end{pmatrix}. \quad (2.322)$$

Let us consider the effects exerted by h_+ for $h_\times = 0$. Then we have the two equations

$$\partial_0^2 X^1 = \frac{1}{2} X^1 \partial_0^2 [h_+ \exp(ik \cdot x)] \quad (2.323)$$

$$\partial_0^2 X^2 = -\frac{1}{2} X^2 \partial_0^2 [h_+ \exp(ik \cdot x)]. \quad (2.324)$$

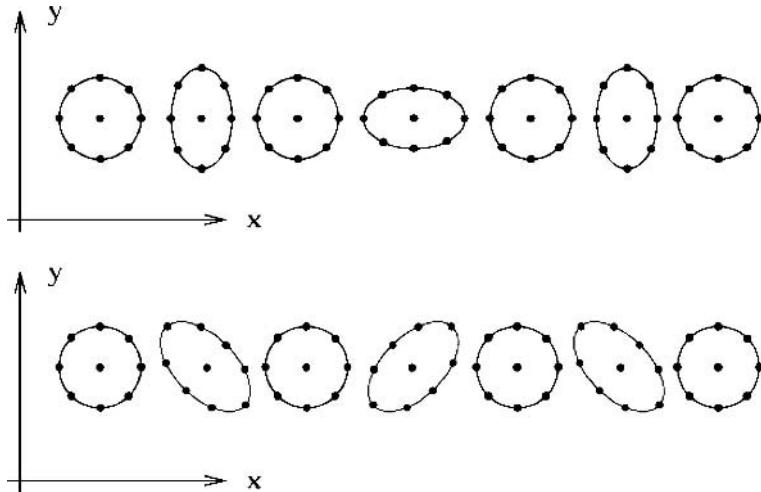


Fig. 2.4. Top: The + mode of gravitational waves. Bottom: The \times mode of gravitational waves. The phases shown are $0, \pi/2, \pi, 3\pi/2, 2\pi$

These can be solved immediately in lowest order as

$$X^1 = \left(1 + \frac{1}{2} h_+ \exp(ik \cdot x) \right) X^1(0) \quad (2.325)$$

$$X^2 = \left(1 - \frac{1}{2} h_+ \exp(ik \cdot x) \right) X^2(0). \quad (2.326)$$

Thus particles initially separated in the x -direction will oscillate in the x -direction, and likewise for those in the y -direction. If we start with a ring of stationary particles in the $x - y$ plane, they will bounce back and forth in the shape of a +, as the wave passes by (Fig. 2.4).

The equivalent analysis for the case where $h_+ = 0$, but $h_x \neq 0$ would yield the solutions

$$X^1 = X^1(0) + \frac{1}{2} X^2(0) h_x \exp(ik \cdot x) \quad (2.327)$$

$$X^2 = X^2(0) + \frac{1}{2} X^1(0) h_x \exp(ik \cdot x). \quad (2.328)$$

In this case, the circle of particles would bounce back and forth in the shape of a \times , as shown in Fig. 2.4. These two quantities measure therefore two independent modes of linear polarization of a gravitational wave, known as the plus and cross polarizations. The sensitivity achieved in 2006 by the GEO600 interferometers is shown in Fig. 2.5. Details for all these experiments are available from their home pages.

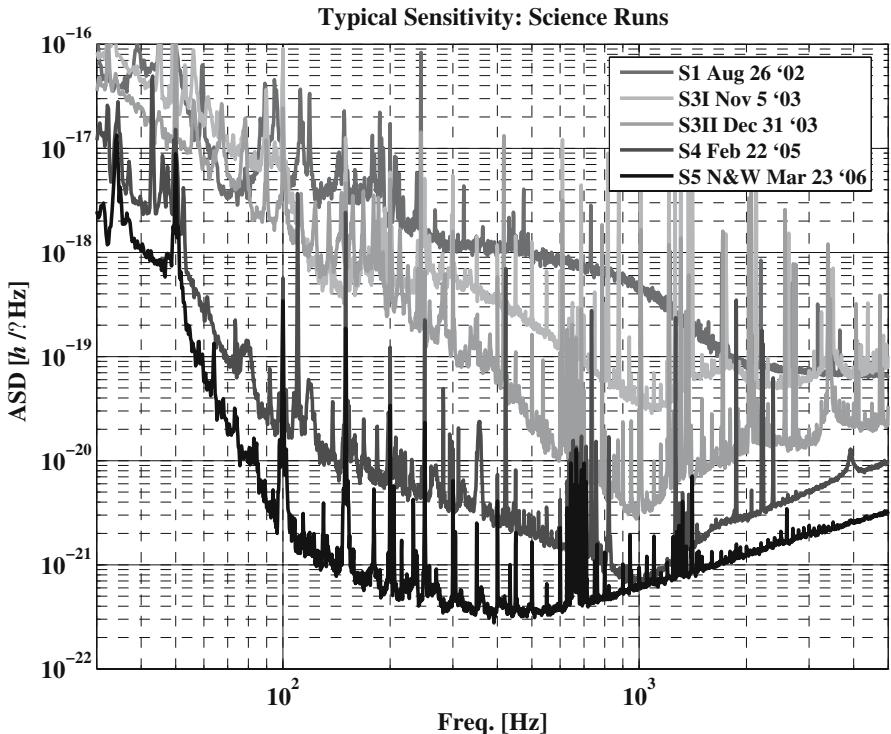


Fig. 2.5. GEO600 sensitivity achieved in various runs. Credit: GEO600 Collaboration

2.7.6 The Quadrupole Formula

Already Albert Einstein [142] found some radiative solutions to the field equations of general relativity and deduced the existence of gravitational waves, which could be interpreted as ripples in the curvature of spacetime propagating with the speed of light. In a subsequent paper [143], he computed the total power emitted in the form of gravitational waves by an isolated Newtonian source, and found that this power depends quadratically on the variations of the quadrupole moment of the source in a celebrated formula known today as **Einstein's quadrupole formula**. The fact that gravitational waves from an isolated source are dominantly quadrupolar is a consequence of the equivalence principle, which implies the conservation (or linear variation in time) of the source's monopole and dipole moments as the consequence of the equations of motion.

Experimental research on gravitational radiation is presently very active with the development of new technologies for bar detectors, and most importantly with the construction of large scale laser interferometric detectors for observation of the waves in the frequency bandwidth between ≈ 10 Hz and ≈ 1000 Hz: the American LIGO detector, the German GEO600 detector (Fig. 2.5) and the Franco-Italian VIRGO detector. These detectors should observe at least one type of source, the

inspiralling compact binary. This source is composed of two neutron stars or black holes spiralling very rapidly around each other in the last rotations preceding their final coalescence. Orbital velocities are much larger than in the binary pulsar for instance, which will coalesce with its companion in few hundreds of millions years.

A useful characterization of slowly moving sources is that their spatial extension is small (and of order ε) as compared to one typical wavelength of the radiation. Indeed the wavelength is given by $\lambda = cP$ where P is a period of motion in the source. But $a \approx vP$ with $v \approx \varepsilon c$, thus $cP \approx a/\varepsilon$ and therefore $a/\lambda \approx \varepsilon$ which is indeed the statement above. Note that since $GM/ac^2 \approx \varepsilon^2$, we also have $GM/\lambda c^2 \approx \varepsilon^3$.

It is convenient for slowly moving sources to introduce an interior domain D_i , sometimes called also the near zone, defined by $D_i = \{(\mathbf{x}, t), |\mathbf{x}| < r_i\}$, where the radius r_i is such that $r_i \approx \varepsilon\lambda$ and $r_i > a$ [78]. The near zone D_i is small with respect to the wavelength of the radiation and covers entirely the source. This is possible only for a slowly moving source. We choose r_i to be strictly larger than a (instead of being a itself) for later convenience.

The near-zone is the domain where one can confidently use the post-Newtonian expansion. The real precision of some post-Newtonian expression will be exactly given, in D_i , by the formal order in ε of the neglected terms. Note that all powers of $1/c$ must be taken into account in finding the magnitude of a term in terms of $\varepsilon \approx 1/c$, including the ones which arise from the temporal gradients $\partial_0 = c^{-1}\partial/\partial t$ (which are really of order ε with respect to the spatial gradients ∂_i). This is clear from the definition of the near zone, where the field is quasistatic and propagation effects are small. Similar arguments are used in the discussion of electromagnetic radiation.

Having clarified the concept of near-zone D_i , it is now necessary to introduce an exterior domain defined by $D_e = \{(\mathbf{x}, t), |\mathbf{x}| > r_e\}$, where the radius r_e is chosen to be strictly between a and r_i , i.e. $a < r_e < r_i$. This choice, which can always be done for slowly moving sources, is in order that the intersection between D_i and D_e (the exterior part of the near zone $D_i \cap D_e$) exists.

Included in the exterior domain D_e , we also consider the so-called exponentially far wave zone, namely the domain $D_w = \{(\mathbf{x}, t), |\mathbf{x}| > r_\omega\}$, where r_ω is a radius such that $r_\omega \approx \lambda e^{\lambda c^2/GM}$. In D_w , where the observer will be located, one can expand the field in powers of the distance of the source, and keep the leading order term in this expansion. The error done in assuming this will be negligible for future astrophysical sources of radiation.

For very relativistic sources like inspiralling compact binaries, the precision given by the Einstein quadrupole formula for the energy in the waves is insufficient. What is required is a relativistic or **post-Newtonian** formalism (involving an expansion when the speed of light goes to infinity) for both the emission and reaction of waves from isolated sources with substantially large internal velocities.

By quadrupole formalism we mean the lowest order formalism, in the post-Newtonian expansion $\varepsilon \rightarrow 0$, for the generation of gravitational radiation from the source, and also for the reaction of the radiation onto the source. Although this may seem to be a little paradoxical, the quadrupole formalism can thus be viewed

as a Newtonian formalism. Relativistic corrections to the formalism are referred to as post-Newtonian corrections accordingly to which post-Newtonian order in the equations of motion of the source (beyond the usual Newtonian acceleration) is needed in order to reduce all accelerations with consistent accuracy. For instance, the first-post-Newtonian (in short 1PN) formalism retains all terms in the radiation field and in the reaction which can be computed consistently using the 1PN equations of motion, which include the corrections ε^2 beyond the Newtonian force (using the standard practice that the n th post-Newtonian order refers to the order ε^{2n}).

The quadrupole formalism for the wave generation expresses the gravitational field h_{ij}^{TT} in a transverse and traceless (TT) coordinate system as

$$h_{ij}^{\text{TT}} = \frac{2G}{c^4 R} \mathcal{P}_{iklm}(\mathbf{n}) \left\{ \frac{d^2 Q_{km}}{dT^2} (T - R/c) + O(\varepsilon) \right\} + O\left(\frac{1}{R^2}\right). \quad (2.329)$$

The coordinate system (\mathbf{x}, T) is centered on the source, with $R = |\mathbf{x}|$ the distance to the source and $\mathbf{n} = \mathbf{x}/R$ the unit direction from the source to the observer. The retarded time at the observer position in D_w is $T - R/c$, and terms of order $1/R^2$ in the distance of the source are neglected (in addition to the post-Newtonian terms of order ε). In front of (2.329) appears

$$\mathcal{P}_{iklm}(\mathbf{n}) = (\delta_{ik} - n_i n_k)(\delta_{jm} - n_j n_m) - \frac{1}{2}(\delta_{ij} - n_i n_j)(\delta_{km} - n_k n_m) \quad (2.330)$$

which is the TT projection operator onto the plane orthogonal to the direction \mathbf{n} . The quadrupole moment of the source takes the familiar Newtonian form

$$Q_{ij}(t) = \int d^3x \varrho(\mathbf{x}, t) \left(x^i x^j - \frac{1}{3} \delta^{ij} \mathbf{x}^2 \right), \quad (2.331)$$

where ϱ denotes the Newtonian mass density of the source. The quadrupole moment (2.331) is taken to be trace-free ($\delta^{ij} Q_{ij} = 0$). This is not important in the waveform (2.329) because at the quadrupolar level the TT operator in front cancels any possible trace in the moment.

The gravitational wave produced by an isolated nonrelativistic object is therefore proportional to the second time derivative of the quadrupole moment of the mass density at the point where the past light cone of the observer intersects the source. In contrast to this, electromagnetic radiation is generated by a time-dependent dipole moment of the charge density. Oscillation of the center-of-mass of an isolated system would violate conservation of momentum. For this, there is no dipolar mode for gravitational waves. Gravitational radiation is typically also much weaker than electromagnetic radiation, since quadrupole moments are in general much smaller than dipole moments.

By differentiating with respect to time the waveform (2.329)–(2.331), squaring the result using standard formulae and integrating over all the directions \mathbf{n} , one obtains the total power in the gravitational waves emitted by the source (see e.g. [2] or [18]),

$$P_{\text{GW}} = \frac{G}{5c^5} \left\{ \sum_{i,j=1}^3 \frac{d^3 Q_{ij}}{dT^3} \frac{d^3 Q_{ij}}{dT^3} + O(\varepsilon^2) \right\}. \quad (2.332)$$

This result is the famous Einstein quadrupole formula [143], which neglects post-Newtonian terms of order ε^2 . It was derived originally by Einstein under the very restrictive assumption that the motion of the source is nongravitational. The quadrupole formula applies also to a gravitationally bound system, for instance a Newtonian binary star system (see Sect. 6.4.4). Many improvements in the derivation of this formula have been obtained later (for this, see Blanchet [78]). The notation P_{GW} stands for the total luminosity of the source in gravitational waves, by analogy with the total luminosity of a star in electromagnetic waves.

The merger of two black holes, or two neutron stars is one of the most extraordinary events in the Universe. Made of pure gravity, the black holes combine to form a single hole, emitting a strong burst of gravitational radiation. Ground-based detectors are currently searching for such bursts from black holes formed in the evolution of binary stars. The space-based LISA detector is being designed to search for such bursts from merging very massive black holes in the centers of galaxies, events that would emit many thousands of solar masses of pure gravitational wave energy over a period of only a few minutes. The black holes spend almost all their time in this first phase of the dynamics, where the stars orbit one another and gradually spiral together. In this first phase (Fig. 2.6), the post-Newtonian approximation

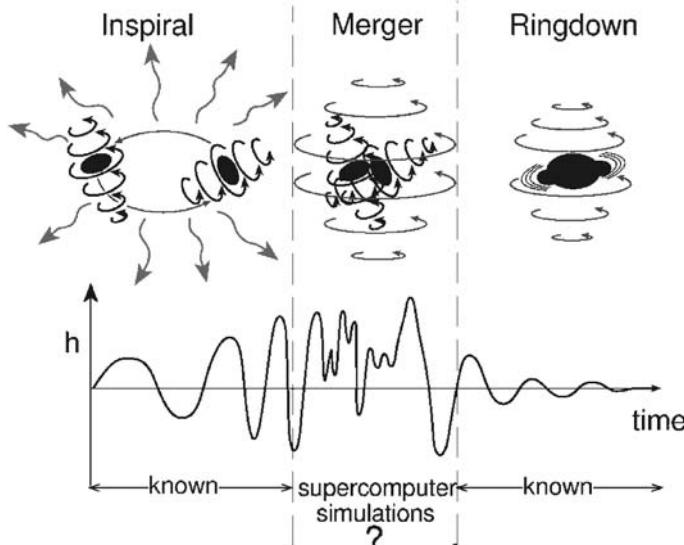


Fig. 2.6. The three phases of black hole merger. The merging of compact objects can be divided up into the inspiral phase, the merger phase, and the ring-down phase. Credit: Figure provided by Kip Thorne

– an asymptotic approximation to general relativity valid for small orbital velocity ($v/c \ll 1$) in gravitationally bound systems – provides a systematic approach to studying the orbital inspiral phase, where orbits shrink and lose eccentricity through the radiation of energy and angular momentum in gravitational waves. The actual merger phase has to be simulated on the computer, since nonlinear effects in the gravitational fields are dominating the time evolution. Finally, the newly formed black hole is relaxing towards a Kerr solution by emission of gravitational waves.

2.8 3+1 Split of Einstein's Equations

In order to numerically solve Einstein's field equations, it is necessary to cast the equations in a form suitable for a computer based treatment. Among the formulations proposed for this purpose, by far the most frequently applied is the canonical **3+1 decomposition** of Arnowitt, Deser and Misner [40] commonly referred to as the ADM formalism. In this approach, spacetime is decomposed into a one-parameter family of three-dimensional space-like hypersurfaces and the Einstein equations are put into the form of an initial value problem. Initial data is provided on one hypersurface in the form of the spatial three-metric and its time derivative and this data is evolved subject to certain constraints and the specification of gauge choices.

It is a known problem, however, that the ADM formalism does not result in a strictly hyperbolic formulation of the Einstein equations and, in combination with its complicated structure, the stability properties of the ensuing finite differencing schemes remain unclear. These difficulties have given rise to the development of modified versions of the ADM formulation in which the Einstein equations are written in a *somewhat more hyperbolic form*. Such modifications of the canonical ADM scheme have been successfully tested, but the optimal 3+1 formulation has yet to be found.

An entirely different approach to the field equations is based on the decomposition of spacetime into families of null surfaces, the characteristic surfaces of the propagation of gravitational radiation [177]. The Einstein field equations are again formulated as an initial value problem and by virtue of a suitable choice of characteristic coordinates one obtains a natural classification of the equations into evolution and hypersurface equations. The characteristic initial value problem was first formulated by Bondi et al. (1962) and Sachs (1962) in order to facilitate a rigorous analysis of gravitational radiation which is properly described at null infinity only. It is a generic drawback of 3+1 formulations that null infinity cannot be included in the numerical grid by means of compactifying spacetime and instead outgoing radiation boundary conditions need to be used at finite radius. Aside from the nonrigorous analysis of gravitational radiation at finite distances these artificial boundary conditions give rise to spurious numerical reflections. A characteristic formulation resolves these problems in a natural way but is itself vulnerable to the formation of caustics in regions of strong curvature.

To perform now the ADM decomposition of a spacetime, we need to introduce a global time function t and some quantities, which develop in time. A globally

hyperbolic spacetime (M, g) possesses these requirements. It can be foliated by Cauchy surfaces Σ_t , which are parametrized by a global time function t .

2.8.1 Induced Spatial Metric and Extrinsic Curvature

Given a manifold M and an open set O , a **family of curves** is said to be a congruence in O , if exactly one curve in the family passes through each point $p \in O$. Due to the uniqueness of the curve at each point, the tangents to a congruence build up into a vector field in O , and every continuous vector field yields a congruence. The congruence is called smooth if the corresponding vector field is smooth.

Subsequently, we consider smooth congruences of time-like geodesics. As we consider only time-like geodesics, we can choose the parametrization by proper time τ , hence, the vector field U^μ of tangents is normalized to unit length, $U^\mu U_\mu = -1$. Now we can define the tensor field $B_{\mu\nu}$

$$B_{\mu\nu} = \nabla_\nu U_\mu , \quad (2.333)$$

with the properties

$$B_{\mu\nu} U^\nu = 0 = B_{\mu\nu} U^\nu . \quad (2.334)$$

Thus, B is said to be purely spatial. The meaning of the introduced tensor field can easily be seen, if we consider a one-parameter subfamily of geodesics in the congruence. Let η^μ be the orthogonal deviation vector from γ_0 for this subfamily. Since the tangent vector field U and the deviation vector field η build a coordinate system, we have $L_U \eta = 0$ and thus

$$U^\nu \nabla_\nu \eta^\mu = \eta^\nu \nabla_\nu U^\mu = B_{\nu}^{\mu} \eta^\nu \quad (2.335)$$

From this, we see that $B_{\mu\nu}$ measures the failure to parallel transport η along the geodesic. $B_{\mu\nu}$ indicates how the geodesics, infinitesimally near to geodesic γ_0 , are twisted and expanded around γ_0 .

Now we are able to split $B_{\mu\nu}$ in several parts: expansion Θ , shear $\sigma_{\mu\nu}$ and twist $\omega_{\mu\nu}$. To perform this decomposition, we first introduce the so-called spatial metric by $\gamma_{\mu\nu} = g_{\mu\nu} + U_\mu U_\nu$. The name “spatial” is justified because $\gamma_v^\mu = g^{\mu\rho} \gamma_{\rho v}$ acts like a projection operator onto the three-dimensional subspace generated by vectors orthogonal to the tangent vector U . Now we are ready to carry out the splitting. The expansion Θ , defined by

$$\Theta = B^{\mu\nu} \gamma_{\mu\nu} = B_\mu^\mu = B \quad (2.336)$$

describes the average convergence or divergence of neighboring geodesics. The shear $\sigma_{\mu\nu}$ defined by

$$\sigma_{\mu\nu} = B_{(\mu\nu)} - \frac{1}{3} \Theta \gamma_{\mu\nu} \quad (2.337)$$

is a symmetric, traceless tensor, and finally the twist $\omega_{\mu\nu} = B_{[\mu\nu]}$ represents the antisymmetric part of $B_{\mu\nu}$. This shows that both, twist and shear, are totally spatial. All together, they build the original tensor

$$B_{\mu\nu} = \frac{1}{3}\Theta\gamma_{\mu\nu} + \sigma_{\mu\nu} + \omega_{\mu\nu}, \quad (2.338)$$

where the first part contains uniquely nonvanishing diagonal elements.

Until now we only have considered congruences of time-like geodesics. In the case of null geodesics there appear difficulties in the construction of the orthogonal deviation vector. For null vectors only two orthogonal, linearly independent vectors can be found and thus the “spatial” metric projects onto a two-dimensional subspace.

So far we just have considered a deviation vector for a one-parameter subfamily of the congruence. The vectors orthogonal to the tangent vector of a time-like geodesic form a three-dimensional, space-like hypersurfaces Σ . The spatial metric γ is the metric induced on this hypersurface with constant proper time τ . The change of γ with proper time can be expressed by the extrinsic curvature $K_{\mu\nu}$, defined by

$$K_{\mu\nu} = -B_{v\mu} = -\nabla_\mu U_v. \quad (2.339)$$

Since the congruence is hypersurface orthogonal, the antisymmetric part of $K_{\mu\nu}$ vanishes, i.e. $\omega_{\mu\nu} = 0$, and thus $K_{\mu\nu}$ is a symmetric tensor. The symmetry enables us to express $K_{\mu\nu}$ in terms of the Lie derivative. Thus we have

$$K_{\mu\nu} = -\frac{1}{2}L_U g_{\mu\nu} = -\frac{1}{2}L_U \gamma_{\mu\nu}. \quad (2.340)$$

The minus sign is a matter of convention. $K_{\mu\nu}$ measures the rate of change of the spatial metric γ as it proceeds along the geodesic congruence. $K_{\mu\nu}$ tells us how $\gamma_{\mu\nu}$, and thus the curvature of the hypersurface Σ , varies in the space-time.

2.8.2 Hypersurface Embedding

Let n^α be the unit orthogonal vector to the hypersurface Σ_t , then the spacetime metric g induces (like above) a spatial metric $\gamma_{\alpha\beta}$ on each Σ_t by

$$\gamma_{\alpha\beta} = g_{\alpha\beta} + n_\alpha n_\beta, \quad (2.341)$$

which is also called the *first fundamental form*. The unit vector n can also be identified with the four-velocity of a class of observers, called *Eulerian observers*, whose worldlines are orthogonal to Σ_t . The one-form \mathbf{n} dual to the vector field n is parallel to the gradient of the scalar field t

$$\mathbf{n} = -N dt. \quad (2.342)$$

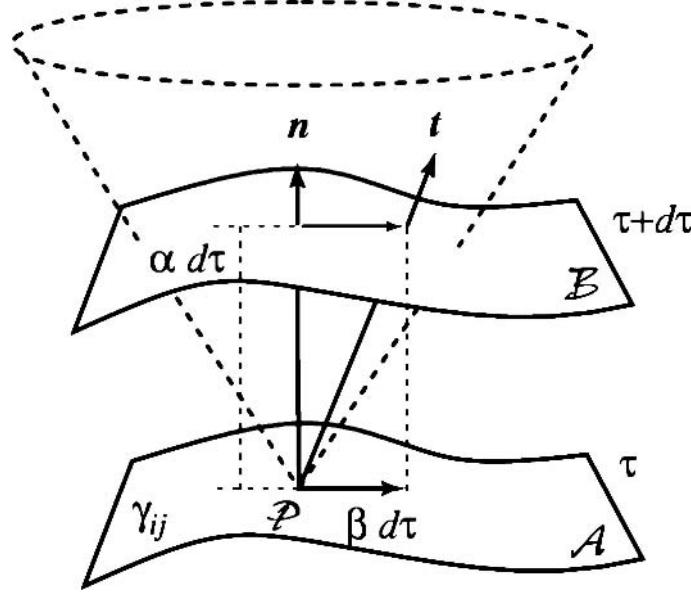


Fig. 2.7. The 3+1 decomposition of spacetime with redshift factor α and shift function β , also called foliation by a family of space-like hypersurfaces Σ_t . The hypersurfaces Σ_t are given as the level surfaces of time t and have a metric γ_{ij}

The proportionality factor N is called *lapse function*. This ensures the normalization condition

$$n \cdot n = \langle \mathbf{n}, \mathbf{n} \rangle = -1 . \quad (2.343)$$

The bending of a hypersurface Σ_t in M is described by the *Weingarten map* or *second fundamental form* which associates with each vector tangent to Σ_t the covariant derivative of the unit normal n along this vector field,

$$\kappa_\beta^\alpha = \gamma_\beta^\mu \nabla_\mu n^\alpha . \quad (2.344)$$

We then define the *extrinsic curvature* of the hypersurface Σ_t as minus the second fundamental form

$$K_{\alpha\beta} = \gamma_\beta^\mu \nabla_\mu n_\alpha . \quad (2.345)$$

Replacing n_α in terms of its gradient leads to

$$K_{\alpha\beta} = \gamma_\beta^\mu (N \nabla_\mu t) . \quad (2.346)$$

Hence

$$K_{\alpha\beta} = -\nabla_\alpha n_\beta - n_\alpha (D_\beta \ln N) . \quad (2.347)$$

The extrinsic curvature is often expressed in terms of the Lie derivative

$$K_{\alpha\beta} = -\frac{1}{2} L_n \gamma_{\alpha\beta}. \quad (2.348)$$

By using integral curves of t^μ to construct a diffeomorphism between Σ_0 and Σ_t , we may view the “flow of time” as changing the spatial metric on a three-dimensional hypersurface from $\gamma(0)$ to $\gamma(t)$. Thus, in a more intuitive image, a globally hyperbolic spacetime (M, g) represents the time development of a Riemannian metric γ on a fixed three-dimensional manifold Σ . Above, we defined the quantity extrinsic curvature K , which turned out to be a measure of change of the induced spatial metric γ on an orthogonal hypersurface Σ , as γ moves along time-like geodesics in a congruence. Up to now the time-like unit tangent vector of a geodesic in the congruence yielded the extrinsic curvature of a hypersurface orthogonal to the geodesics. But now we have only the foliation with Cauchy surfaces Σ_t and a unit vector field n^μ orthogonal to these Cauchy surfaces. If we suppose ξ^μ to be the tangent vectors of geodesics of a congruence on the Cauchy surface Σ_t , both vector fields, ξ and n , coincide on the hypersurface Σ_t . And thus, the extrinsic curvature can be expressed in terms of the normal vector field n^μ by

$$K_{\mu\nu} = -\nabla_\mu \xi_\nu = -\gamma_\mu^\rho \nabla_\rho \xi_\nu = -\frac{1}{2} L_n \gamma_{\mu\nu}. \quad (2.349)$$

Hence, we do not have the requirement of geodesics to construct the notion of extrinsic curvature.

As the vector field denotes the flow of time and the unit normal vector are related through $n = (1, -\beta)/\alpha$, we get for the time derivative

$$\begin{aligned} \partial_t \gamma_{\mu\nu} &= L_t \gamma_{\mu\nu} = \alpha L_n \gamma_{\mu\nu} + L_\beta \gamma_{\mu\nu} \\ &= -2\alpha K_{\mu\nu} + L_\beta \gamma_{\mu\nu} \\ &= -2\alpha K_{\mu\nu} + D_\mu \beta_\nu + D_\nu \beta_\mu. \end{aligned} \quad (2.350)$$

2.8.3 Split of Affine Connection and Curvature

The above splitting of the manifold into a foliation of three-surfaces will induce a corresponding splitting of the affine connection. For this we define an orthonormal tetrad field given by

$$e_0 = \frac{1}{\alpha} (\partial_t - \beta^i \partial_i) \quad (2.351)$$

$$e_i = \bar{e}_i, \quad (2.352)$$

where \bar{e}_i is an orthonormal triad on the hypersurface Σ . The corresponding dual one-forms are given by

$$\Theta^0 = \alpha dt \quad (2.353)$$

$$\Theta^i = \bar{\Theta}^i + \beta^i dt, \quad (2.354)$$

where $\bar{\Theta}^i(\bar{e}_k) = \delta_k^i$. The corresponding spacetime metric is then

$$ds^2 = [-\alpha^2 + \beta^i \beta_i] dt^2 + 2\beta_i dt dx^i + \gamma_{ik} dx^i dx^k \quad (2.355)$$

which is often written as

$$ds^2 = \alpha^2 dt^2 + \gamma_{ik}(dx^i + \beta^i dt)(dx^k + \beta^k dt). \quad (2.356)$$

α is called the **lapse function** or redshift factor and β^i are the **shift vectors**.

The connection of a spacetime is given by six one-forms ω_a^b which can be decomposed into

$$\omega_j^0 = \omega_j^0(e_0)\Theta^0 + \omega_j^0(\bar{e}_i)\Theta^i \quad (2.357)$$

$$\begin{aligned} \omega_j^i &= \omega_j^i(\bar{e}_m)\Theta^m + \omega_j^i(e_0)\Theta^0 \\ &= \omega_j^i(e_0)\Theta^0 + \omega_j^i(\bar{e}_m) \left[\bar{\Theta}^m + \frac{1}{\alpha}\beta^m\Theta^0 \right] \\ &= \omega_j^i(\bar{e}_m)\Theta^m + \left[\omega_j^i(e_0) + \frac{1}{\alpha}\omega_j^i(\beta) \right] \Theta^0. \end{aligned} \quad (2.358)$$

The coefficient $\omega_i^0(e_0)$ follows from the first structure equation

$$d\Theta^0 = \alpha_{,i}\bar{\Theta}^i \wedge dt = \nabla_i \ln \alpha \Theta^i \wedge \Theta^0 = -\nabla_i \ln \alpha \Theta^0 \wedge \Theta^i, \quad (2.359)$$

i.e.

$$\omega_i^0(e_0) = \nabla_i \ln \alpha. \quad (2.360)$$

From the definition of the extrinsic curvature, $K_{ij} = -n_{i|j} = -e_{0|j}$, and the definition of the covariant derivative

$$\nabla_X e_0 = \omega_0^i(X)e_i \quad (2.361)$$

we find

$$\omega_i^0(e_j) = -K_{ij}. \quad (2.362)$$

With the definition of $\partial_t \bar{\Theta}^i = c_j^i \bar{\Theta}^j$ we can calculate the exterior derivatives

$$\begin{aligned} d\Theta^i &= d\bar{\Theta}^i + dt \wedge \partial_t \bar{\Theta}^i + d\beta^i \wedge dt \\ &= -\bar{\omega}_j^i \wedge \bar{\Theta}^j + \frac{1}{\alpha}\Theta^0 \wedge [c_j^i \bar{\Theta}^j - d\beta^i]. \end{aligned} \quad (2.363)$$

By comparing this with the first structure equation

$$\begin{aligned} d\Theta^i &= -\omega_0^i \wedge \Theta^0 - \omega_j^i \wedge \Theta^j \\ &= -\omega_i^0(e_m)\Theta^m \wedge \Theta^0 - \omega_j^i(e_0)\Theta^0 \wedge \Theta^j - \omega_j^i(e_m)\Theta^m \wedge \Theta^j, \end{aligned} \quad (2.364)$$

we find

$$\begin{aligned} & [K_{ij} + \omega_{ij}(e_0) + \frac{1}{\alpha} \bar{\omega}_{ij}(\beta)] \Theta^j \wedge \Theta^0 \\ &= -\frac{1}{\alpha} c_{ij} \Theta^j \wedge \Theta^0 + \frac{1}{\alpha} d\beta^i \wedge \Theta^0 + \frac{1}{\alpha} \bar{\omega}_{ij}(\beta) \Theta^j \wedge \Theta^0. \end{aligned} \quad (2.365)$$

Since $d\beta^i + \bar{\omega}^i_j(e_m) \beta^m \Theta^j = D\beta^i$ is the covariant derivative in the hypersurface, we gain the identity

$$\omega_{ij}(e_0) + K_{ij} = \frac{1}{\alpha} [\beta_{i|j} - c_{ij} - \bar{\omega}_{ij}(\beta)] \quad (2.366)$$

with its solution for the symmetric part

$$K_{ij} = \frac{1}{\alpha} [\beta_{(i|j)} - c_{(ij)}] \quad (2.367)$$

and for the antisymmetric part

$\omega_{ij}(e_0) = \frac{1}{\alpha} [\beta_{[i|j]} - c_{[ij]} - \bar{\omega}_{ij}(\beta)] .$

(2.368)

In total, we have derived a closed expression for the connection forms

$$\omega^0_j = (\nabla_j \ln \alpha) \Theta^0 - K_{ij} \Theta^j \quad (2.369)$$

$$\omega^i_j = \bar{\omega}^i_j + H^i_j \Theta^0. \quad (2.370)$$

The connection of spacetime is given by the connection of the hypersurface, the gravitational force $\nabla \ln \alpha$, the extrinsic curvature K_{ij} and the antisymmetric gravitomagnetic field \mathbf{H} defined as

$$H_{ij} = \frac{1}{\alpha} [\beta_{[i|j]} - c_{[ij]}] . \quad (2.371)$$

Split of the Curvature

The Gauss–Codazzi equations are a collection of equations which relate the four-dimensional Riemann tensor R_{abcd} , Ricci tensor R_{ab} and Ricci scalar R to their projection onto a three-dimensional hypersurface embedded within four-dimensional spacetime. For this we consider the decomposition of the curvature two-form. For the curvature on the three-surface Σ we find the Gauss form (for a derivation, see Appendix C)

$$\Omega^i_j|_\Sigma = \bar{\Omega}^i_j + K^i_s K_{jt} \bar{\Theta}^s \wedge \bar{\Theta}^t . \quad (2.372)$$

In components, this means for the Riemann tensor of the hypersurface

$${}^{(3)}R_{ijkl} = h_i^p h_j^q h_k^r h_m^s R_{pqrs} - K_{ik} K_{jm} + K_{im} K_{jk} , \quad (2.373)$$

where $h_{ab} = g_{ab} + n_a n_b$ is the projection tensor with $n_a n^a = -1$. The Codazzi–Mainardi equation is then given by

$$\Omega^0_i(e_j, e_0) = \frac{1}{\alpha} D_j(\alpha_{,i}) + \frac{1}{\alpha} (\partial_t - L_\beta) K_{ij} + K_{jm} K_i^m . \quad (2.374)$$

2.8.4 Split of Einstein's Equations

For Einstein's equations we only need the split of the Ricci tensor

$$R_{00} = \Omega^i_0(e_i, e_0) = \Omega^0_i(e_i, e_0) \quad (2.375)$$

$$R_{0i} = \Omega^j_0(e_j, e_i) = \Omega^0_j(e_j, e_i) = \Omega^0_j(\bar{e}_j, \bar{e}_i) \quad (2.376)$$

$$R_{ij} = \Omega^0_i(e_0, e_j) + \Omega^m_i(e_m, e_j). \quad (2.377)$$

Poisson Equation

For the Ricci component R_{00} we need the trace of equation (C.12). Thereby

$$\gamma^{im} \partial_t K_{im} = \partial_t \text{Tr}(\mathbf{K}) - K_{im} \partial_t \gamma^{im} = \partial_t \text{Tr}(\mathbf{K}) + K^{im} \partial_t \gamma_{im}. \quad (2.378)$$

As usual, Tr denotes the trace of a matrix. Similarly, we find for the Lie derivative

$$\begin{aligned} \gamma^{im} L_\beta K_{im} &= L_\beta \text{Tr}(\mathbf{K}) - K_{im} L_\beta \gamma^{im} \\ &= L_\beta \text{Tr}(\mathbf{K}) + K^{im} L_\beta \gamma_{im} \\ &= L_\beta \text{Tr}(\mathbf{K}) + K^{im} \left[2\alpha K_{im} + \partial_t \gamma_{im} \right]. \end{aligned} \quad (2.379)$$

In this way, the time component of Einstein's equations provides the generalization of the Poisson equation

$$\frac{1}{\alpha} \bar{\Delta} \alpha + \frac{1}{\alpha} (\partial_t - L_\beta) \text{Tr}(\mathbf{K}) - \text{Tr}(\mathbf{K}^2) = 8\pi G \left(E + \frac{1}{2} \text{Tr}(T) \right). \quad (2.380)$$

$\bar{\Delta}$ is the Laplacian on the hypersurface. Since in the Newtonian limit $\alpha \simeq 1 + \Phi$, this equation is the natural generalization of the Poisson equation. It determines the redshift factor α in terms of the matter distribution and some nonlinearities due to the embedding of the hypersurface into spacetime.

Constraint Equations

The most widely used approach for analyzing solutions of Einstein's gravitational field equations is the initial value formulation, and the key to the initial value formulation is the set of Einstein constraint equations. The study of the constraint equations and their solutions plays a major role in the search to understand the degrees of freedom of the gravitational field, in the search for a theory of gravity consistent with the quantum principle, and in the search for physically realistic models of astrophysical and cosmological gravitational systems. During the past few years, there has been significant progress in our understanding of the Einstein constraints.

The relativistic Poisson equation still includes a time dependence. Einstein's equations in fact contain four equations which are time-independent. For this we consider the G_{00} component

$$\begin{aligned}
G_{00} &= R_{00} - \frac{1}{2} \eta_{00} R \\
&= R_{00} + \frac{1}{2} (R_0^0 + R_i^i) \\
&= R_{00} - \frac{1}{2} R_{00} + \frac{1}{2} R_{ii} = \frac{1}{2} (R_{00} + R_{ii}) \\
&= \frac{1}{2} (\Omega_i^0(e_i, e_0) + \Omega_i^0(e_0, e_i) + \Omega_i^k(e_k, e_i)) \\
&= \frac{1}{2} \Omega_i^k(e_k, e_i). \tag{2.381}
\end{aligned}$$

G_{00} and R_{0i} will only depend on the curvature forms that are defined on the three-space

$$G_{00} = \frac{1}{2} \Omega_i^k(\bar{e}_k, \bar{e}_i) \tag{2.382}$$

$$R_{0i} = \Omega_j^0(\bar{e}_j, \bar{e}_i). \tag{2.383}$$

We now insert the above results into G_{00} to obtain

$$\begin{aligned}
G_{00} &= \frac{1}{2} [\bar{\Omega}_i^k + K_s^k K_{it} \bar{\Theta}^s \wedge \bar{\Theta}^t](\bar{e}_k, \bar{e}_i) \\
&= \frac{1}{2} \left[\sum_i \bar{R}_{ii} + K_s^k K_{it} (\bar{\Theta}^s(\bar{e}_k) \bar{\Theta}^t(\bar{e}_i) - \bar{\Theta}^s(\bar{e}_i) \bar{\Theta}^t(\bar{e}_k)) \right] \\
&= \frac{1}{2} [\bar{R} + K_s^k K_{it} (\delta_k^s \delta_i^t - \delta_i^s \delta_k^t)] \\
&= \frac{1}{2} [\bar{R} + K_k^k K_{tt} - K_i^k K_{ik}] \\
&= \frac{1}{2} [\bar{R} + [Tr(K)]^2 - Tr(K^2)], \tag{2.384}
\end{aligned}$$

and similarly for R_{0i}

$$\begin{aligned}
R_{0i} &= -D_K j_s \wedge \bar{\Theta}^s(\bar{e}_j, \bar{e}_i) \\
&= -D_m K_{js} \bar{\Theta}^m \wedge \bar{\Theta}^s(\bar{e}_j, \bar{e}_i) \\
&= -D_m K_{js} (\delta_j^m \delta_i^s - \delta_i^m \delta_j^s) \\
&= -D_j K_{ji} + D_i K_{jj} \\
&= D_i K_j^j - D_j K_i^j. \tag{2.385}
\end{aligned}$$

These **constraints equations** can be summarized

$$\bar{R} + [Tr(K)]^2 - Tr(K^2) = 16\pi G T_{(00)} \tag{2.386}$$

$$D_i K_j^j - D_j K_i^j = -8\pi G T_{(0i)}. \tag{2.387}$$

They have to be satisfied for all times and on any time-slice. The first step in the building of a spacetime solution of Einstein's gravitational field equations via the

initial value formulation is finding a solution of the Einstein constraint equations. One useful step towards this goal is learning how to glue together known solutions of the constraint equations. We will not discuss these topics in detail here.

Evolution Equations

The remaining six equations determine the time evolution of the three-metric of the slices. The spatial part of the Ricci tensor follows from the decomposition of equation (C.12) and equation (C.3)

$$R_{ij} = \bar{R}_{ij} - \frac{1}{\alpha} \alpha_{|ij} + Tr(\mathbf{K}) K_{ij} - 2\mathbf{K}_{ij}^2 - \frac{1}{\alpha} (\partial_t - L_\beta) K_{ij} . \quad (2.388)$$

This is the fundamental dynamical equation in Einstein's theory. The dynamical equations are then often written as a pair of equations for the metric of the three-surface γ_{ij} and for the extrinsic curvature

$$(\partial_t - L_\beta) \gamma_{ij} = -2\alpha K_{ij} \quad (2.389)$$

$$(\partial_t - L_\beta) K_{ij} = \alpha \bar{R}_{ij}(\gamma) - \alpha_{|ij} + \alpha Tr(\mathbf{K}) K_{ij} - 2\alpha \mathbf{K}_{ij}^2 - 4\pi G \alpha (2S_{ij} - \gamma_{ij} T) . \quad (2.390)$$

They have to be solved together with the four constraints equations

$$\bar{R} + [Tr(K)]^2 - Tr(K^2) = 16\pi G E \quad (2.391)$$

$$\bar{D}_i K_j^j - \bar{D}_j K_i^j = -8\pi G S_i . \quad (2.392)$$

In the above equations the energy density $E = T_{(00)}$, the momentum density S_i and the spatial stress density S_{ij} are functions of the nongravitational fields and the metric as well as their derivatives.

Now, we have a system of second-order differential equations for the spatial metric γ_{ij} on a three-dimensional manifold Σ . Given the initial data $(\gamma_{ij}(0), K_{ij}(0))$ on Σ , the evolution of γ_{ij} constructs a globally hyperbolic spacetime (M, g) and Σ is a Cauchy hypersurface of (M, g) , on which the initial data are induced.

2.8.5 Black Hole Simulations and Gravitational Waves

The fundamental importance of obtaining stable accurate simulations of black holes arises from the fact that the inspiral and merger of binary black holes is considered one of the most promising sources of detectable gravitational waves.

Among the various difficulties one faces in black hole simulations we highlight the two most pronounced ones (for more details, see e.g. [322]):

- A generic difficulty one faces in general relativity arises from the coordinate invariance of the theory. For numerical purposes this means that the physical scenario is independent of the coordinate system used for the simulation. The

numerical properties of a code, however, will depend very sensitively on the choice of coordinates, i.e. gauge conditions. It is a highly nontrivial, and as yet unsolved, task to find gauge conditions suitable for the simulation of strongly dynamic black hole systems.

- A second major difficulty is a particular feature of black holes and can be illustrated easily in the case of a single black hole. The extrinsic curvature takes on increasingly large values in the vicinity of the black hole. In fact it becomes infinite at the black hole center. It is known that the center of a stationary single black hole represents a true physical singularity, which cannot be described by traditional numerical techniques. When a computer encounters such a singularity in an evolution, it will simply crash, because it cannot handle infinite numbers. In order to circumvent this problem, one can proceed along two different lines.
 - Using so-called singularity avoiding slicing one ensures that a numerical simulation slows down near a singularity and does not reach such a point in a finite time.
 - A technique called Black Hole Excision simply removes a finite region around the black hole singularity from the computational domain.

The main motivation for using black hole excision arises from Cosmic Censorship which states that a singularity is always surrounded by an event horizon, a boundary of finite extension which causally disconnects the interior with the singularity from the outside world. Whatever goes on inside the event horizon will under no circumstances affect the physics in the outer region. In particular one may therefore remove a region inside the event horizon from the numerical evolution without observing any changes in the outside evolution.

Problems

2.1. Gauss' Theorem: Since n -forms on a manifold M of dimension n are proportional to the volume form, the volume integration over a n -form gives a mapping to the real numbers. For any $(n - 1)$ -form ω on M , $d\omega$ is an n -form, which can be integrated over M , while ω itself can be integrated over ∂M . Stokes' theorem tells us then that [11]

$$\int_M d\omega = \int_{\partial M} \omega. \quad (2.393)$$

Let us consider a one-form $A = A_n dx^n$. Its Hodge dual defines then a $(n - 1)$ -form ω

$$\omega = *A, \quad (2.394)$$

or in components

$$\omega_{i_1 \dots i_{n-1}} = A^a \eta_{ai_1 \dots i_{n-1}}, \quad (2.395)$$

where η is the Levi-Civita n -form on M . We can reconstruct A from ω by means of

$$A = (-1)^{s+n-1} * \ast A = (-1)^{s+n-1} * \omega. \quad (2.396)$$

s equals +1 for Euclidean signatures and -1 for Lorentzian signatures.

Prove that the exterior derivative of ω is proportional to the divergence of A , i.e.

$$d\omega = (\nabla_a A^a) \sqrt{|g|} d^n x. \quad (2.397)$$

Similarly, ω itself can be represented as a volume-form on ∂M

$$\omega = (n_a A^a) \sqrt{|\gamma|} d^{n-1} x, \quad (2.398)$$

where n^a is the unit normal to the boundary ∂M and γ is the intrinsic metric of ∂M . Putting both formulae together, Stokes' theorem relates the divergence of a vector field A to its flux through the boundary

$$\int_M (\nabla_a A^a) \sqrt{|g|} d^n x = \int_{\partial M} (n_a A^a) \sqrt{|\gamma|} d^{n-1} y. \quad (2.399)$$

The unit normal n should be chosen outward-pointing if the boundary is space-like, and inward-pointing for time-like boundaries.

2.2. Spin Connection of the Three-Sphere: The metric of a three-sphere is given in the coordinates (ψ, θ, ϕ)

$$ds^2 = a^2 \left[d\psi^2 + \sin^2 \psi (d\theta^2 + \sin^2 \theta d\phi^2) \right]. \quad (2.400)$$

Find orthonormal frames e_i and dual frames \varTheta^i such that the metric tensor becomes δ_{ik} . Compute the component of the spin connection by solving Cartan's first structure equation. Compute the components of the Riemann tensor by solving the second Cartan structure equation. Calculate the Christoffels in a coordinate frame and the corresponding Riemann tensors. Compare your results with the Cartan method. Confirm that the three-sphere is a maximally symmetric space.

2.3. Light Deflection in Weak Gravitational Fields: Use the metric for weak gravitational fields

$$ds^2 = -[1 + 2\Phi(\mathbf{x})] dt^2 + [1 - 2\gamma\Phi(\mathbf{x})] d\mathbf{x}^2 \quad (2.401)$$

to derive the light deflection in the gravitational potential $\Phi(\mathbf{x})$. The Robertson parameter γ can be used to test Einstein's gravity [419].

2.4. Shapiro Time Delay in Solar System: In general relativity, the Shapiro effect, or gravitational time delay, is one of the four classic Solar System tests of general relativity. It says that a radar beam (or light beam) which passes near a massive object, as it travels from some observer's location to a target and returns to the

observer, takes slightly longer to make the round trip (as measured by the observer) than it would if the object were not present. More generally, the travel time of any signal moving at the local speed of light can be affected by the gravitational field in regions of spacetime through which it travels. In general relativity (and in most other gravitation theories), the local speed of light is a constant of nature, but the time-delay effect implies that the effective global speed of light is path-dependent.

The gravitational time delay, now known as the Shapiro effect, was first predicted by Einstein. It is proven by measuring the time delay of radar and radio beams sent between Earth and Mars. As the beams pass closer and closer to the Sun, a delay in the transit time is measured. This delay is caused by the gravitational force of the Sun. The time-delay effect was first noticed in 1964, by Irwin I. Shapiro [364] (see also [419]). Shapiro proposed an observational test of his prediction: bounce radar beams off the surface of Venus and Mercury, and measure the round trip travel time. When the Earth, Sun, and Venus are most favorably aligned, Shapiro showed that the expected time delay, due to the presence of the Sun, of a radar signal travelling from the Earth to Venus and back, would be about 200 milliseconds, well within the limitations of 1960s era technology.

Discuss the Shapiro time delay in the weak field geometry (2.401) and derive the classic formula for the time delay in its round-trip travel time, given by

$$\Delta\tau_S = \frac{2R_S}{c} \frac{1+\gamma}{2} \ln \left[\frac{(r_{\oplus} + \mathbf{n} \cdot \mathbf{x}_{\oplus})(r_E - \mathbf{n} \cdot \mathbf{x}_E)}{d^2} \right], \quad (2.402)$$

where \mathbf{x}_E (\mathbf{x}_{\oplus}) are the vectors, and r_E (r_{\oplus}) are the distances from the Sun to the emitter (Earth), respectively. R_S is the Schwarzschild radius of the Sun.

2.5. Gravitational Wave Amplitude for a Binary System: Let us consider two stars of mass M in a circular orbit at a distance R from their common center-of-mass. Treat the motion of the stars in their Newtonian approximation, i.e. as Kepler orbits. Calculate the wave amplitude $h_{ij}^{TT}(t, \mathbf{x})$ of a gravitational wave emitted by this circular binary system, with binary period P_b , if it is located at distance d from a wave detector. Consider in particular a binary neutron star system in the Virgo cluster with a separation $R \simeq 100$ km, i.e. for $d \simeq 16$ Mpc, and compare the amplitude with the present sensitivity of LIGO.

3 Matter Models for Compact Objects

Besides gravity, the description of various forms of matter is essential for compact objects. We discuss two matter models, the hydrodynamic model and the Boltzmann model.

3.1 General Relativistic Hydrodynamics

Many applications for compact objects are based on a hydrodynamical description of matter: the internal structure of white dwarfs and neutron stars is based on the hydrostatic approximation, and accretion onto compact objects in general requires a time-dependent treatment of gas dynamics. We can define a **perfect fluid** such that in local comoving coordinates the fluid is isotropic. Assuming that the spacetime is Minkowskian, the energy–momentum tensor of the fluid is given by

$$T^{tt} = \varrho, \quad T^{xx} = T^{yy} = T^{zz} = P, \quad (3.1)$$

where ϱ is the total proper energy density and P the pressure. When each fluid element has a spatial velocity v^i with respect to some fixed lab frame, the expression of the energy–momentum tensor is obtained via a Lorentz boost

$$T^{\mu\nu} = (\varrho + P) u^\mu u^\nu + P \eta^{\mu\nu}. \quad (3.2)$$

Here, u^μ is the fluid four-velocity, satisfying $u^\mu u_\mu = -1$. The equations for conservation of energy and momentum can be written as $T^{\mu\nu}_{;\nu} = 0$ in Minkowski spacetime. In order to extend this expression to curved spacetime we only need to replace the Minkowskian metric η by the general Lorentz metric of the spacetime and partial derivatives with covariant ones. Thus, in a general curved spacetime, the stress energy tensor for a perfect fluid (plasma) is given by

$$T^{\mu\nu} = (\varrho + P) u^\mu u^\nu + P g^{\mu\nu}. \quad (3.3)$$

In the strong gravity regime, pressure and stresses are typically so large that we cannot assume that the fluid is incompressible. In addition, the pressure contributions to the stress tensor can be of the same order as those from the energy density. This makes GR plasmas behave very differently from the type of plasmas that we encounter in daily life, where the stress energy tensors are dominated by their rest-mass density.

3.1.1 Relativistic Plasma Equations

The general relativistic hydrodynamic equations consist of the local conservation laws of the stress–energy tensor $T^{\alpha\beta}$ (the Bianchi identities) and of the matter current density $J^\alpha = \varrho_0 u^\alpha$ (the continuity equation)

$$\nabla_\alpha T^{\alpha\beta} = 0 \quad (3.4)$$

$$\nabla_\alpha J^\alpha = 0. \quad (3.5)$$

In distinction to the energy density ϱ , we denote the rest-mass energy density as ϱ_0 . The above expression for the stress–energy tensor can be extended to a nonperfect plasma as follows (see, e.g. Straumann [18])

$$T^{\alpha\beta} = \varrho u^\alpha u^\beta + (P - \zeta \Theta) h^{\alpha\beta} - 2\eta \sigma^{\alpha\beta} + q^\alpha u^\beta + q^\beta u^\alpha, \quad (3.6)$$

where $h^{\alpha\beta}$ is the spatial projection tensor $h^{\alpha\beta} = g^{\alpha\beta} + u^\alpha u^\beta$. In addition, η and ζ are the shear and bulk viscosities. The expansion Θ , describing the divergence or convergence of the fluid worldlines, is defined as $\Theta = \nabla_\alpha u^\alpha$. The symmetric, trace-free, spatial shear tensor σ is defined by

$$\sigma^{\alpha\beta} = \frac{1}{2} ((\nabla_\varrho u^\alpha) h^{\varrho\beta} + (\nabla_\varrho u^\beta) h^{\varrho\alpha}) - \frac{1}{3} \Theta h^{\alpha\beta}. \quad (3.7)$$

Finally, q^α is the heat energy flux vector, which is space-like, $u_\alpha q^\alpha = 0$.

In the following, we mainly consider ideal gas dynamics, i.e. we will neglect nonadiabatic effects, such as viscosity or heat transfer, assuming the stress–energy tensor to be that of a **perfect fluid**

$$T^{\alpha\beta} = \varrho u^\alpha u^\beta + P h^{\alpha\beta}. \quad (3.8)$$

In a local chart, the previous conservation equations read

$$\partial_\alpha \left[\sqrt{-g} J^\alpha \right] = 0 \quad (3.9)$$

$$\partial_\alpha \left[\sqrt{-g} T^{\alpha\beta} \right] = -\sqrt{-g} \Gamma_{\varrho\alpha}^\beta T^{\varrho\alpha}. \quad (3.10)$$

In order to close the system, the equations of motion and the continuity equation must be supplemented with an equation of state (EoS) relating some fundamental thermodynamical quantities. In general, the EoS takes the form $P = P(\varrho_0, \epsilon)$. Traditionally, most of the approaches for numerical integrations of the general relativistic hydrodynamic equations have adopted space-like foliations of the spacetime, within the 3+1 formulation. Recently, however, covariant forms of these equations, well suited for advanced numerical methods, have also been developed.

Special Relativistic Hydrodynamics

In the framework of special relativity, the motion of an ideal fluid is governed by particle number conservation and energy–momentum conservation. In the lab frame

of reference, these two conservation equations can be written in closed divergence form (Martí and Müller [271])

$$\boxed{\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}^i}{\partial x^i} = 0.} \quad (3.11)$$

The five-dimensional state vector $\mathbf{U} = (D, S_i, E)^T$, $i = 1, 2, 3$, consists of the relativistic density D , the momentum density three-vector \mathbf{S} and the total energy density τ with pressure P . The transformation between the rest frame quantities ϱ , the specific enthalpy h , pressure P and velocity \mathbf{v} are given by

$$D = W\varrho_0 \quad (3.12)$$

$$\mathbf{S} = \varrho_0 W^2 h \mathbf{v} \quad (3.13)$$

$$\tau = \varrho_0 W^2 h - P - D, \quad (3.14)$$

where $W = 1/\sqrt{1 - \mathbf{v}^2}$ is the Lorentz factor and $h = 1 + e/\varrho_0 + P/\varrho_0$ the relativistic specific enthalpy. The corresponding flux vectors are given by

$$\mathbf{F}^i = \left(Dv^i, S_j v^i + P\delta_j^i, (\tau + P)v^i \right). \quad (3.15)$$

The state of the relativistic plasma is therefore given either in terms of the five-dimensional state vector $\mathbf{U} = \mathbf{U}(\mathbf{P})$, or in terms of the primitive variables $\mathbf{P} = (\varrho, v_1, v_2, v_3, P)^T$. While the expression for the state vector \mathbf{U} in terms of the primitive variables \mathbf{P} is trivial, the inverse relation involves the calculation of the Lorentz factor

$$\varrho_0 = D/W \quad (3.16)$$

$$\mathbf{v} = \mathbf{S}/(E + P) \quad (3.17)$$

$$P = DWh - E. \quad (3.18)$$

The Lorentz factor can be expressed in terms of the pressure

$$\frac{1}{W^2(P)} = 1 - \frac{\mathbf{S}^2}{(E + P)^2}. \quad (3.19)$$

For given D , \mathbf{S} and E , one can derive from the above relations an implicit expression for P

$$f(P) = Dh(P, \tau)W(P) - E - P = 0, \quad (3.20)$$

where $\tau = 1/\varrho$ denotes the specific proper volume, which is related to the enthalpy variation

$$dh|_s = \tau dP. \quad (3.21)$$

This equation must be solved for all grid points in order to recover the pressure from the values of the state vector \mathbf{U} .

To numerically solve equations (3.11), each spatial dimension is discretized into cells. The time-dependent evolution can then be expressed in the semidiscrete form

$$\begin{aligned} \frac{d\mathbf{U}_{i,j,k}}{dt} = & -\frac{\mathbf{F}_{i+1/2,j,k}^x - \mathbf{F}_{i-1/2,j,k}^x}{\Delta x} \\ & -\frac{\mathbf{F}_{i,j+1/2,k}^y - \mathbf{F}_{i,j-1/2,k}^y}{\Delta y} \\ & -\frac{\mathbf{F}_{i,j,k+1/2}^z - \mathbf{F}_{i,j,k-1/2}^z}{\Delta z}, \end{aligned} \quad (3.22)$$

where $\mathbf{F}_{i+1/2,j,k}^x$, etc. are the fluxes at the cell interface. The subscripts $i \pm 1/2$, $j \pm 1/2$, and $k \pm 1/2$, refer to cell edges. In order to achieve high accuracy in time, the time integration is done using a higher order Runge–Kutta scheme (Shu and Osher [369]). Such a scheme combines the first-order Euler steps and involves prediction and correction. Third-order accuracy can be obtained with the scheme

$$\mathbf{U}^{(1)} = \mathbf{U}^n + \Delta t L(\mathbf{U}^n) \quad (3.23)$$

$$\mathbf{U}^{(2)} = \frac{3}{4}\mathbf{U}^n + \frac{1}{4}\mathbf{U}^{(1)} + \frac{1}{4}\Delta t L(\mathbf{U}^{(1)}) \quad (3.24)$$

$$\mathbf{U}^{n+1} = \frac{1}{3}\mathbf{U}^n + \frac{2}{3}\mathbf{U}^{(2)} + \frac{2}{3}\Delta t L(\mathbf{U}^{(2)}). \quad (3.25)$$

The operator $L(\mathbf{U})$ is the right-hand side of equation (3.22), \mathbf{U}^{n+1} is the final value of the state vector after advancing one time-step from \mathbf{U}^n . Fluxes at the cell edges are calculated by using the technique of Riemann solvers and suitable cell reconstruction methods (Marti and Müller [271]). Modern computer codes have to solve these equations using adaptive mesh refinement (AMR) on parallel computers.

3+1 Split of the Plasma Equations in Curved SpaceTime

The special relativistic hydrodynamic equations can be easily generalized to any curved spacetime by using the covariant formulation (3.10). In the *Valencia* formulation of Papadopoulos and Font [325], the spatial velocity components v^i of the four-velocity, u^i , as measured by an Eulerian observer at rest in the space-like hypersurface Σ_t with normal \mathbf{n}

$$v^i = \frac{u^i}{\alpha u^t} + \frac{\beta^i}{\alpha} \quad (3.26)$$

together with the rest-frame density and internal energy, ϱ_0 and ϵ , provide a unique description of the state of the fluid at a given time and are taken as the **primitive variables**. They constitute a vector in a five-dimensional space, $\mathbf{P} = (\varrho_0, v^i, \epsilon)^T$. The initial value problem for hydrodynamical equations is defined in terms of another

vector in the same fluid state space, namely the **conserved variables**, \mathbf{U} , individually denoted $\mathbf{U} = (D, S_i, \tau)^T$, where

$$D = -J^\mu n_\mu = \varrho_0 W \quad (3.27)$$

$$S_j = -h_{\mu j} T^{\mu\nu} n_\nu = \varrho_0 h W^2 v_j \quad (3.28)$$

$$\tau = T^{\mu\nu} n_\nu n_\nu + J^\mu n_\mu = \varrho_0 h W^2 - P - D. \quad (3.29)$$

$n_\mu = (-\alpha, 0, 0, 0)$ is the unitary vector normal to the slice, $W = -u^\mu n_\mu = \alpha u^t$ the Lorentz factor, and $h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$ denotes the projection operator, $u_j = W v_j$. $h = 1 + \epsilon + P/\varrho_0$ is the enthalpy.

It is now important that the GR hydrodynamic equations can also be cast into conservation law form. For this we begin with $(T^\mu_\nu)_{;\mu} = 0$ and perform the following manipulations

$$\begin{aligned} (g_{\delta\nu} T^{\mu\delta})_{;\mu} &= T^{\mu\delta} g_{\delta\nu;\mu} + g_{\delta\nu} T^{\mu\delta}_{;\mu} \\ &= g_{\delta\nu} \left[\frac{1}{\sqrt{-g}} (\sqrt{-g} T^{\mu\delta})_{,\mu} + \Gamma_{\mu\lambda}^\delta T^{\mu\lambda} \right] \\ &= g_{\delta\nu} \left[\frac{1}{\sqrt{-g}} \left(\sqrt{-g} g^{\varrho\delta} T_\varrho^\mu \right)_{,\mu} + \Gamma_{\mu\lambda}^\delta T^{\mu\lambda} \right] \\ &= g_{\delta\nu} \left[g_{\cdot\mu}^{\varrho\delta} T_\varrho^\mu + g^{\varrho\delta} \frac{1}{\sqrt{-g}} \left(\sqrt{-g} T_\varrho^\mu \right)_{,\mu} + \Gamma_{\mu\lambda}^\delta T^{\mu\lambda} \right]. \end{aligned} \quad (3.30)$$

From this expression we derive the fundamental relation

$$\frac{1}{\sqrt{-g}} (\sqrt{-g} T_\nu^\mu)_{,\mu} = g_{\varrho\nu,\mu} T^{\mu\varrho} - \Gamma_{\mu\lambda}^\delta T^{\mu\lambda}. \quad (3.31)$$

The left-hand side is now a true divergence, and on the right-hand side we find the corresponding source terms. These terms vanish in a Minkowskian spacetime based on Cartesian coordinates, and this form leads to the expression (3.11).

Since $u^t = W/\alpha$, where $W = 1/\sqrt{1 - v_i v^i}$ is the Lorentz factor for the three-velocity \mathbf{v} , we can factor out the redshift factor α from the above equations by using the expression for the determinant, $\sqrt{-g} = \alpha \sqrt{\gamma}$, where γ_{ij} is the metric of the hypersurface Σ_t in the 3+1 split of spacetime. With these definitions, the equations of general relativistic hydrodynamics take the standard conservation law form,

$$\frac{1}{\sqrt{-g}} \left[\frac{\partial [\sqrt{\gamma} \mathbf{U}]}{\partial t} + \frac{\partial [\sqrt{-g} \mathbf{F}^i]}{\partial x^i} \right] = \boldsymbol{\delta}.$$

(3.32)

The flux vectors \mathbf{F}^i and the source terms $\boldsymbol{\delta}$ (which depend only on the metric, its derivatives and the undifferentiated stress energy tensor), are given by ($i = 1, 2, 3$)

$$\mathbf{F}^i = \left(D (v^i - \beta^i/\alpha), S_j (v^i - \beta^i/\alpha) + P \delta_j^i, \tau (v^i - \beta^i/\alpha) + Pv^i \right) \quad (3.33)$$

and

$$\delta = \left(0, T^{\mu\nu} \left[\partial_\mu g_{\nu j} - \Gamma_{\mu\nu}^\delta g_{\delta j} \right], \alpha \left[T^{\varrho t} \partial_\varrho \ln \alpha - T^{\varrho\sigma} \Gamma_{\varrho\sigma}^t \right] \right). \quad (3.34)$$

In analogy to special relativity, the flux vector can also be written as

$$\mathbf{F}^i = \left(Du^i/W, S_j u^i/W + P \delta_j^i, \tau u^i/W + Pv^i \right). \quad (3.35)$$

It is therefore quite convenient to work with the modified three-velocity $V^i = u^i/W$ instead of v^i . In contrast to special relativity, curved coordinates add a source term to the divergence equation, apart from the fact that the divergence is now formulated in a covariant manner. This source term is essentially the result from the Christoffel symbols in (3.10).

The state of the fluid is uniquely described using either vector of variables, i.e. either \mathbf{U} or \mathbf{P} , and each one can be obtained from the other via their definitions and the use of the normalization condition for the four-velocity. The local characteristic structure of the above system of equations was presented by Papadopoulos and Font [326], where the formulation proved well suited for the numerical implementation of high-resolution shock capturing schemes (HRSC schemes).

The interaction between matter and radiation fields, present in different levels of complexity in all astrophysical systems, is described by the equations of radiation hydrodynamics. The Newtonian framework is mostly based on the method of flux limited diffusion. Pons et al. [330] discuss a hyperbolic formulation of the radiative transfer equations, paying particular attention to the closure relations and to extend HRSC schemes to those equations. General relativistic formulations of radiative transfer in curved spacetimes are still underdeveloped.

3.1.2 On Numerics of Hydrodynamics

Any system of equations presented in the previous section can be solved numerically by replacing the partial derivatives by finite differences on a discrete numerical grid, and then advancing the solution in time via some time-marching algorithm. Hence, specification of the state vector \mathbf{U} on an initial hypersurface, together with a suitable choice of EoS, followed by a recovery of the primitive variables, leads to the computation of the fluxes and source terms. Through this procedure the first time derivative of the data is obtained, which then leads to the formal propagation of the solution forward in time, with a time-step constrained by the Courant–Friedrichs–Lewy (CFL) condition.

The hydrodynamic equations (either in Newtonian physics or in general relativity) constitute a nonlinear hyperbolic system and, hence, smooth initial data can transform into discontinuous data (the crossing of characteristics in the case of shocks) in a finite time during the evolution. As a consequence, classical finite difference schemes present important deficiencies when dealing with such systems. Typically, first-order accurate schemes are much too dissipative across discontinuities (excessive smearing) and second-order (or higher) schemes produce spurious

oscillations near discontinuities, which do not disappear as the grid is refined. To avoid these effects, standard finite difference schemes have been conveniently modified in various ways to ensure high-order, oscillation-free accurate representations of discontinuous solutions.

In all conservative formulations discussed in the literature, the time update of a given numerical algorithm is applied to the conserved quantities \mathbf{U} . After this update the vector of primitive quantities \mathbf{P} must be re-evaluated, as those are needed in the Riemann solver. The relation between the two sets of variables is, in general, not in closed form and, hence, the recovery of the primitive variables is done using a root-finding procedure, typically a Newton–Raphson algorithm, as described in the special relativistic treatment. This feature, distinctive of the equations of (special and) general relativistic hydrodynamics – it does not exist in the Newtonian limit – may lead in some cases to accuracy losses in regions of low density and small speeds, apart from being computationally inefficient. In particular, for the covariant formulations, there exists an analytic method to determine the primitive variables, which is, however, computationally very expensive since it involves many extra variables and solving a quartic polynomial. Therefore, iterative methods are still preferred.

The equations of viscous hydrodynamics, the Navier–Stokes–Fourier equations, have been formulated in relativity in terms of causal dissipative relativistic fluids (see the Living Reviews article by Marti and Müller [271], and Peitz and Appl [327]). These extended fluid theories, however, remain unexplored, numerically, in astrophysical systems. The reason may be the lack of an appropriate formulation well-suited for numerical studies. Work in this direction was done by Peitz and Appl [327] who provided a 3+1 coordinate-free representation of different types of dissipative relativistic fluid theories which possess, in principle, the potentiality of being well adapted to numerical applications. The inclusion of magnetic fields and the development of formulations for the MHD equations will be discussed in the section on accretion.

Artificial Viscosity

The idea of modifying the hydrodynamic equations by introducing artificial viscosity terms to damp the amplitude of spurious oscillations near discontinuities was originally proposed by von Neumann and Richtmyer (1950) in the context of the (classical) Euler equations. The basic idea is to introduce a purely artificial dissipative mechanism whose form and strength are such that the shock transition becomes smooth, extending over a small number of intervals Δx of the space variable.

The main advantage of the artificial viscosity approach is its simplicity, which results in high computational efficiency. Experience has shown, however, that this procedure is both problem dependent and inaccurate for ultrarelativistic flows. Furthermore, the artificial viscosity approach has the inherent ambiguity of finding the appropriate form for Q that introduces the necessary amount of dissipation to reduce the spurious oscillations and, at the same time, avoids introducing excessive smear-

ing in the discontinuities. In many instances both properties are difficult to achieve simultaneously.

High-Resolution Shock-Capturing (HRSC) Upwind Schemes

In finite difference schemes, convergence properties under grid refinement must be enforced to ensure that the numerical results are correct (i.e. if a scheme with an order of accuracy is used, the global error of the numerical solution has to tend to zero as the cell width tends to zero). For hyperbolic systems of conservation laws, schemes written in conservation form are preferred since, according to the Lax–Wendroff theorem, they guarantee that the convergence, if it exists, is to one of the so-called weak solutions of the original system of equations. Such weak solutions are generalized solutions that satisfy the integral form of the conservation system. They are classical solutions (continuous and differentiable) in regions where they are continuous and have a finite number of discontinuities.

The Lax–Wendroff theorem cited above does not establish whether the method converges. To guarantee convergence, some form of stability is required, as Lax first proposed for linear problems. Along this direction, the notion of total-variation stability has proven very successful, although powerful results have only been obtained for scalar conservation laws. The total variation of a solution at time t^n , $TV(u^n)$, is defined as

$$TV(u^n) = \sum_{j=0}^{\infty} |u_{j+1}^n - u_j^n|. \quad (3.36)$$

A numerical scheme is said to be TV stable if $TV(u^n)$ is bounded for all Δt at any time for each initial data. In the case of nonlinear, scalar conservation laws it can be proved that TV stability is a sufficient condition for convergence, as long as the numerical schemes are written in conservation form and have consistent numerical flux functions. Current research has focused on the development of high-resolution numerical schemes in conservation form satisfying the condition of TV stability, such as the so-called total variation diminishing (TVD) schemes (Harten 1984).

Let us now consider the specific system of hydrodynamic equations as formulated in (3.32), and let us consider a single computational cell of our discrete spacetime. Let Ω be a region (simply connected) of a given four-dimensional manifold, bounded by a closed three-dimensional surface $\partial\Omega$. We further take the three-surface $\partial\Omega$ as the standard-oriented hyperparallelepiped made up of two space-like surfaces ($\Sigma_t, \Sigma_{t+\Delta t}$) plus time-like surfaces ($\Sigma_x, \Sigma_{x+\Delta x}$) that join the two temporal slices together. By integrating system (3.32) over a domain Ω of a given spacetime, the variation in time of the state vector \mathbf{U} within Ω is given – keeping apart the source terms – by the fluxes \mathbf{F}^i through the boundary $\partial\Omega$. The integral form of system (3.32) is

$$\int_{\Omega} \frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} \mathbf{U}] d\Omega + \int_{\Omega} \frac{1}{\sqrt{-g}} \partial_i [\sqrt{-g} \mathbf{F}^i] d\Omega = \int_{\Omega} \mathbf{s} d\Omega. \quad (3.37)$$

Hence the time variation of the mean value of \mathbf{U}

$$\bar{U} = \frac{1}{\Delta V} \int_{x^1}^{x^1 + \Delta x^1} \int_{x^2}^{x^2 + \Delta x^2} \int_{x^3}^{x^3 + \Delta x^3} \sqrt{\gamma} \mathbf{U} dx^1 dx^2 dx^3 \quad (3.38)$$

within the spatial volume

$$\Delta V = \int_{x^1}^{x^1 + \Delta x^1} \int_{x^2}^{x^2 + \Delta x^2} \int_{x^3}^{x^3 + \Delta x^3} \sqrt{\gamma} dx^1 dx^2 dx^3 \quad (3.39)$$

can be obtained by applying the Gauss theorem

$$\begin{aligned} (\bar{U} \Delta V)_{t+\Delta t} - (\bar{U} \Delta V)_t = & \\ & + \left(\int_{\Sigma_{x^1}} \sqrt{-g} \mathbf{F}^1 dx^t dx^2 dx^3 - \int_{\Sigma_{x^1 + \Delta x^1}} \sqrt{-g} \mathbf{F}^1 dx^t dx^2 dx^3 \right) \\ & + \left(\int_{\Sigma_{x^2}} \sqrt{-g} \mathbf{F}^2 dt dx^1 dx^3 - \int_{\Sigma_{x^2 + \Delta x^2}} \sqrt{-g} \mathbf{F}^2 dt dx^1 dx^3 \right) \\ & + \left(\int_{\Sigma_{x^3}} \sqrt{-g} \mathbf{F}^3 dt dx^1 dx^2 - \int_{\Sigma_{x^3 + \Delta x^3}} \sqrt{-g} \mathbf{F}^3 dt dx^1 dx^2 \right) \\ & + \int_{\Omega} \mathcal{S} d\Omega. \end{aligned} \quad (3.40)$$

In order to update the solution in time, the volume and surface integrals on the right-hand side have to be evaluated. HRSC schemes rely on the calculation of the fluxes through the volume boundaries by solving local Riemann problems combined with suitable cell reconstruction schemes. An exhaustive discussion on Riemann solvers can be found in the textbook by Toro [396]. The equations in integral form are then advanced, e.g. with a second-order Runge–Kutta method [369].

Even in curved spacetime we can apply Minkowskian Riemann solvers, since, according to the equivalence principle, physical laws in a local inertial frame of a curved spacetime have the same form as in special relativity. The details of this construction can be found in Pons et al. [330].

3.2 The Boltzmann Equation in GR

For many applications in the physics of compact objects, the hydrodynamical description is not appropriate. When photons and neutrinos are involved, we need the general relativistic version of the Boltzmann equation. In this section, we give a short outline of the basic elements.

3.2.1 The Geodesics Spray on the Cotangent Bundle

Particles move on geodesics in the manifold – except for short interactions. Their motion follows from the Lagrangian \mathcal{L} on $\mathcal{T}\mathcal{M}_4$ (see Sect. 2.4.5)

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2} g(\dot{x}, \dot{x}). \quad (3.41)$$

With the canonical transformation between the tangent bundle and the cotangent bundle

$$\mathcal{T}\mathcal{M} \rightarrow \mathcal{T}^*\mathcal{M} : (x^\alpha, \dot{x}^\alpha) \rightarrow (x^\alpha, p_\alpha) , \quad p_\alpha \equiv g_{\alpha\beta} \dot{x}^\beta , \quad (3.42)$$

we arrive at the Hamiltonian

$$\mathcal{H}(x, p) = \frac{1}{2} g^{\alpha\beta} p_\alpha p_\beta . \quad (3.43)$$

On the cotangent bundle, the motion is determined in terms of the Hamiltonian vector field, which is given in natural coordinates

$$X_E = p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma_{\nu\lambda}^\alpha p^\nu p^\lambda \frac{\partial}{\partial p^\alpha} .$$

(3.44)

The integral curves satisfy the canonical equations

$$\frac{dx^\alpha}{d\lambda} = p^\alpha \quad (3.45)$$

$$\frac{dp^\alpha}{d\lambda} = -\Gamma_{\beta\gamma}^\alpha p^\beta p^\gamma . \quad (3.46)$$

We now introduce the 7D *one-particle phase space* by means of the set of all points in the cotangent bundle

$$P_m = \{(x, p) \quad , \quad g(x)(p, p) = -m^2\} . \quad (3.47)$$

Its fiber $P_m(x)$ is often called the *mass shell*. The restriction of X_E will be denoted by X_g . On $\mathcal{T}^*\mathcal{M}$ we also use the volume form

$$\begin{aligned} \Omega_g &= (-g) \left[dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 \right] \wedge \left[dp^0 \wedge dp^1 \wedge dp^2 \wedge dp^3 \right] \\ &= (-g) d^4x \wedge d^4p . \end{aligned}$$

(3.48)

This volume form induces a volume form μ_m on the phase space P_m , which is invariant under X_g , i.e.

$$L_{X_g} \mu_m = 0 .$$

(3.49)

The volume form μ_m can be constructed as follows. From the Hamiltonian mechanics we know that

$$\Omega_g = -d\mathcal{L} \wedge \mu_m , \quad (3.50)$$

where a natural choice for this volume is given by

$$\mu_m = \frac{|g|}{-p_0} d^4x d^3p , \quad (3.51)$$

because

$$\begin{aligned} -d\mathcal{L} \wedge \mu_m &= [-g_{\alpha\beta} p^\alpha dp^\beta + \dots] \wedge \mu_m \\ &= (-g) d^4x \wedge g_{\alpha 0} p^\alpha dp^0 \wedge \frac{d^3p}{|p_0|} = \Omega_g . \end{aligned} \quad (3.52)$$

Hence, μ_m is the product of two volumes, one defined on the base manifold and the second one on the mass shell

$$\mu_m = \eta \wedge \Pi_m . \quad (3.53)$$

η is the natural volume form on the manifold and Π_m the natural volume form on the mass shell, as discussed in special relativity,

$$\Pi_m = \sqrt{-g} \frac{d^3p}{|p_0|} . \quad (3.54)$$

We can now define a six-form on the phase space defined as follows

$$\omega_m = i_{X_g} \mu_m , \quad (3.55)$$

which is closed, $d\omega_m = 0$, since

$$d\omega_m = (d \cdot i_{X_g}) \mu_m = L_{X_g} \mu_m = 0 , \quad (3.56)$$

with $d\mu_m = 0$ and $L_{X_g} \mu_m = 0$. By using the above decomposition for μ_m , we obtain

$$\omega_m = (i_{X_g} \eta) \wedge \Pi_m + \eta \wedge i_{X_g} \Pi_m . \quad (3.57)$$

Let Σ now be a time-section of the phase space in the sense that Σ defines a space-like hypersurface on the phase space. Then the number of particles on this time section is proportional to its volume

$$N[\Sigma] = \int_{\Sigma} f(x, p) \omega_m . \quad (3.58)$$

$f(x, p)$ is called one-particle distribution function on the phase space. This definition corresponds to the covariant generalization of the Newtonian one-particle distribution

$$dN(t, \mathbf{x}; \mathbf{p}) = f(t, \mathbf{x}; \mathbf{p}) d^3x d^3p . \quad (3.59)$$

dN denotes the number of particles measured in a volume d^3x at time t and at the spatial position \mathbf{x} with momenta between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$.

3.2.2 Particle Number Current and Energy–Momentum Tensor

The particle number current is then given by the first moment of the one-particle distribution function

$$N^\alpha = \int_{P_m(x)} p^\alpha f(x, p) \Pi_m , \quad (3.60)$$

and the energy–momentum tensor by means of the second moment in momentum space

$$T^{\alpha\beta} = \int_{P_m(x)} p^\alpha p^\beta f(x, p) \Pi_m . \quad (3.61)$$

We can now show that particle number conservation follows from

$$N^\alpha_{;\alpha} = \int_{P_m(x)} (L_{X_g} f) \Pi_m . \quad (3.62)$$

First of all, one finds that

$$d(f\omega_m) = (L_{X_g} f) \mu_m , \quad (3.63)$$

since

$$df \wedge \omega_m = df \wedge i_{X_g} \mu_m = (i_{X_g} df) \wedge \mu_m = (L_{X_g} f) \mu_m . \quad (3.64)$$

Now let us consider a domain D in the 7D phase space P_m with boundary ∂D . Then we find from the expression for the volume element (3.57), setting $i_X \eta = X^\alpha \Sigma_\alpha$,

$$\int_{\partial D} f\omega_m = \int_{\partial D_3} \Sigma_\alpha \int_{P_m(x)} p^\alpha f \Pi_m = \int_{\partial D_3} \Sigma_\alpha N^\alpha = \int_{D_3} (\nabla \cdot N) \eta . \quad (3.65)$$

On the other hand, we can also write directly by applying Stokes' theorem

$$\int_{\partial D} f\omega_m = \int_D d(f\omega_m) = \int_D (L_{X_g} f) \mu_m = \int_{D_3} \eta \int_{P_m(x)} (L_{X_g} f) \Pi_m . \quad (3.66)$$

Since D_3 is arbitrary, we indeed obtain the expression for the number current conservation.

Similarly, one can prove the conservation of the energy–momentum tensor

$$T^{\alpha\beta}_{;\beta} = \int_{P_m(x)} p^\alpha (L_{X_g} f) \Pi_m . \quad (3.67)$$

One has to replace the vector field N^α by a vector field $t^\alpha = T^{\alpha\beta} u_\beta$, where u^β denotes a geodesic flow at x .

3.2.3 The Relativistic Boltzmann Equation

When particles are not suffering from collisions, the one-particle distribution is conserved along a geodesic spray

$$L_{X_g} f = 0. \quad (3.68)$$

This is known as the *collisionless Boltzmann equation*, or Liouville equation.

The proof is very simple. For this we consider a cylinder \mathcal{C} on the 8D cotangent bundle, which is generated by the geodesic spray. The time-section Σ is then mapped to a time-section Σ_t by the geodesics, spanned by the time interval $[0, t]$ (Fig. 3.1). We now consider the volume integral over this cylinder

$$\int_{\mathcal{C}} d(f\omega_m) = \int_{\mathcal{C}} (L_{X_g} f) \mu_m = \int_{\partial\mathcal{C}} f \omega_m. \quad (3.69)$$

Here, we used the identity (3.63). The integral over the mantle vanishes (particles are not leaving the side surface which is parallel to the geodesics), and the two contributions from Σ and Σ_t cancel each other due to conservation of particles. This has the consequence that the particle current is really conserved, $N_{;\alpha}^\alpha = 0$.

Including collisions, the Boltzmann equation is symbolically written as

$$L_{X_g} f = Q[f], \quad (3.70)$$

where $Q[f]$ denotes the collision integral. This involves the microphysics taking place in the interactions, e.g. in the form of the transition matrix elements for two-body collisions.

With this form of the general Boltzmann equation, we now get an expression for the conservation of the energy-momentum

$$T^{\alpha\beta}_{;\beta} = Q^\alpha, \quad (3.71)$$

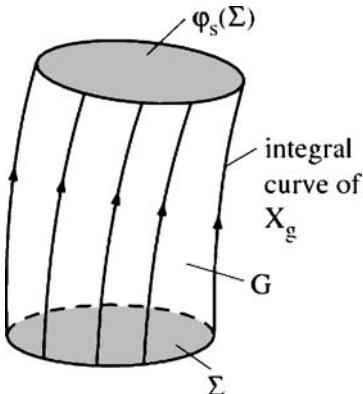


Fig. 3.1. A geodesic spray defines a mapping of a time-section Σ and generates a cylindrical volume G in the cotangent bundle of spacetime

with

$$Q^\alpha = \int_{P_m(x)} p^\alpha Q[f] \Pi_m . \quad (3.72)$$

The conservation of the energy-momentum depends on the first moment of the collision integral which does not vanish in general.

3.2.4 Liouville Operator in 3+1 Split

For applications, it is suitable to transform to Eulerian observers (also called FIDOs, fiducial observers)

$$\left\{ \mathbf{e}_0 = \frac{1}{\alpha} (\partial_t - \beta^i e_i), \mathbf{e}_i \right\} , \quad g(\mathbf{e}_a, \mathbf{e}_b) = \eta_{ab} . \quad (3.73)$$

The four-momenta, p^a , and distribution function $f = f(t, \mathbf{x}; \nu, p^A)$ are also given in this system. This leads to the expression for the Liouville operator ($a = 0, 1, 2, 3$; $A = 1, 2; I = 0, A; h \equiv 1 \equiv c$)

$$L_X f = p^a \mathbf{e}_a(f) - \omega_a^I(p) p^a \frac{\partial f}{\partial p^I} \quad (3.74)$$

or, for photons with frequency ν

$$L_X f = \frac{\nu}{\alpha} \partial_t f + \left(p^a - \frac{\nu}{\alpha} \beta^a \right) \partial_a f - \omega_a^0(p) p^a \frac{\partial f}{\partial \nu} - \omega_m^A(p) p^m \frac{\partial f}{\partial p^A} . \quad (3.75)$$

ω_b^a are the connection forms with respect to Eulerian observers.

As we have seen in Sect. 2.8.3, the connection can be decomposed for stationary spacetimes according to the scheme of the 3+1 split

$$\omega_i^0(\mathbf{e}_0) = \nabla_i \ln \alpha \quad (3.76)$$

$$\omega_i^0(\mathbf{e}_k) = -K_{ik} , \quad K_{ik} = \frac{1}{2\alpha} [\beta_{i|k} + \beta_{k|i}] \quad (3.77)$$

$$\omega_k^i(\mathbf{e}_0) = -\frac{1}{\alpha} \omega_k^i(\mathbf{B}) + H_k^i \quad (3.78)$$

$$\omega_k^i(\mathbf{e}_m) : \text{connection of } \Sigma_t . \quad (3.79)$$

With this, we get the decomposition of the Liouville operator (for photons, where $\mathbf{p} \equiv \nu \mathbf{n}$ is the photon momentum for the Eulerian observer)

$$\begin{aligned} L_X f &= \frac{\nu}{\alpha} \frac{\partial f}{\partial t} + \nu \left(n^i - \frac{1}{\alpha} \beta^i \right) \frac{\partial f}{\partial x^i} \\ &\quad - \nu \left[(\mathbf{n} \cdot \nabla) \ln \alpha - K_{ik} n^i n^k \right] \nu \frac{\partial f}{\partial \nu} \\ &\quad - \nu \left[\nabla^i \ln \alpha - H_k^i n^k + \omega_k^i (\mathbf{n} - \mathbf{B}/\alpha) n^k \right] \frac{\partial f}{\partial n^i} . \end{aligned} \quad (3.80)$$

The first two terms are the standard expressions, except for the change of the frequency by the redshift factor. Gravitational forces, given by $\nabla \ln \alpha$, and the extrinsic curvature K_{ik} lead to a frequency shift in the distribution function, while gravitational forces, the gravitomagnetic field H_{ik} and the connection of the hypersurface imply directional changes. In curved spacetime, the Boltzmann equation is a real differential equation in all seven dimensions.

3.2.5 Transformation into the Local Rest Frame

Since the collision integral for the Boltzmann equation is only known in the local rest frame (LRFM) of matter, and not in the Eulerian frame, we transform the Liouville operator into the LRFM by means of a local Lorentz transformation $\Lambda(x)$

$$\hat{\mathbf{e}}_a = \Lambda^b{}_a \mathbf{e}_b . \quad (3.81)$$

This means for the momenta

$$\hat{p}^a = \bar{\Lambda}^a{}_b p^b . \quad (3.82)$$

From this we recover the transport equation in the LRFM

$$\Lambda^a{}_b \hat{p}^b \mathbf{e}_a(f) - \hat{\omega}^I{}_a(\hat{p}) \hat{p}^a \frac{\partial f}{\partial \hat{p}^I} = \hat{Q} ,$$

(3.83)

where $f = f(x, \hat{p})$ is the distribution function in the LRFM. The connection forms follow from the local transformation of connection forms

$$\hat{\omega}^a{}_b(\hat{e}_s) = \Lambda^a{}_m \left[\bar{\Lambda}^r{}_b \omega_r^m(e_n) + d \bar{\Lambda}^a{}_b(e_n) \right] \bar{\Lambda}^n{}_s . \quad (3.84)$$

The photon momentum \hat{p} is given in terms of the frequency ν_0 and two direction cosines, or in terms of the unit vector \mathbf{n}_0 with $\mathbf{n}_0 = \mathbf{n}_0(\mu, \chi)$, in the local rest frame of matter

$$\hat{p} = \frac{\hbar \nu_0}{c} (1, \mathbf{n}_0) . \quad (3.85)$$

Instead of the phase-space distribution, one uses in astronomy the specific intensity

$$\hat{I}_\nu = \hat{I}_\nu(x, \mathbf{n}_0) \propto \nu^3 \hat{f} . \quad (3.86)$$

For photons, we may express the Lorentz transformation between the LRFM and Eulerian observers by means of a Doppler factor and aberration

$$\nu = \gamma \nu_0 (1 + \mathbf{n}_0 \cdot \mathbf{v}/c) \quad (3.87)$$

$$\mathbf{n} = \frac{\nu_0}{\nu} \left(\mathbf{n}_0 + \gamma \frac{\mathbf{v}}{c} \left[1 + \frac{\gamma}{1+\gamma} \frac{\mathbf{n}_0 \cdot \mathbf{v}}{c} \right] \right) . \quad (3.88)$$

With the definition of the Ricci rotation coefficients $\hat{\Gamma}_{bc}^a \equiv \hat{\omega}_b^a(\hat{e}_c)$, we can define a three-vector with components ($I = 0, 2, 3$)

$$m^I \equiv \hat{\Gamma}_{ab}^I \hat{p}^a \hat{p}^b = v_0^2 \left[\hat{\Gamma}_{00}^I + (\hat{\Gamma}_{0k}^I + \hat{\Gamma}_{k0}^I) n_0^k + \hat{\Gamma}_{km}^I n_0^k n_0^m \right]. \quad (3.89)$$

Therefore, the photon transport equation assumes finally the form

$$\begin{aligned} & \frac{\gamma}{\alpha} (1 + \mathbf{v} \cdot \mathbf{n}_0) (\partial_t - \mathbf{f} \cdot \nabla) \hat{I}_v + (\mathbf{n}_0 \cdot \mathbf{e}) \hat{I}_v \\ & - m^0 \frac{\partial \hat{I}_v}{\partial v_0} - m^A \frac{\partial \hat{I}_v}{\partial n_0^A} = v_0^2 \hat{Q}. \end{aligned} \quad (3.90)$$

For thermodynamic equilibrium between radiation and matter, the collision operator \hat{Q} is given in terms of the Planck function B_v , the absorption opacity $\kappa^{(\text{abs})}$ and the scattering operator Q_{sc} by means of

$$v_0^2 \hat{Q} = \kappa^{(\text{abs})} \varrho (B_v - I_v) + Q_{\text{sc}}. \quad (3.91)$$

Problems

3.1. Special Relativistic Hydrodynamical Equations: Give expressions for the special relativistic hydrodynamical equations (3.11) in cylindrical coordinates.

3.2. Give expressions for the special relativistic hydrodynamical equations (3.11) in spherical coordinates.

3.3. Relativistic Blast Waves: Relativistic blast waves are believed to produce afterglow emission of gamma-ray bursts. The explosion is driven by a central trigger that creates a relativistic shell with Lorentz factor $\Gamma_{\text{ej}} \sim 10^2\text{--}10^3$. This ejecta expands and drives a forward shock (FS) into the external medium, and a reverse shock (RS) propagates inside the ejecta. Such a standard explosion picture has four regions: 1 – external medium; 2 – shocked external medium; 3 – shocked ejecta; and 4 – unshocked ejecta. Regions 2 and 3 are separated by a contact discontinuity.

Consider a relativistic spherically symmetric blast wave. A relativistic ideal gas with four-velocity u^α has stress–energy tensor $T^{\alpha\beta} = (e + p)u^\alpha u^\beta + g^{\alpha\beta}p$ and mass flux $J^\alpha = \varrho u^\alpha$. Here $g_{\alpha\beta} = \text{diag}(-1, 1, r^2, r^2 \sin^2 \theta)$ is Minkowski metric in spherical coordinates (t, r, θ, ϕ) , ϱ is rest-mass density, e is energy density (including rest energy), and p is pressure; all these quantities are defined in the rest frame of the gas.

Show that a relativistic shock is described by three jump conditions that express the continuity of mass-flux J^α and stress–energy $T^{\alpha\beta}$ in the frame of the shock front (see e.g. Landau and Lifshitz 1959),

$$W_2 \beta_2 \varrho_2 = W_1 \beta_1 \varrho_1 \quad (3.92)$$

$$W_2^2 \beta_2 (e_2 + p_2) = W_1^2 \beta_1 \varrho_1 c^2 \quad (3.93)$$

$$W_2^2 \beta_2^2 (e_2 + p_2) + p_2 = W_1^2 \beta_1^2 \varrho_1 c^2. \quad (3.94)$$

Here subscripts 1 and 2 refer to the preshock (cold) and postshock (hot) medium, β_1 and β_2 are the gas velocities relative to the shock front, and we assumed $p_1 = 0$ and $e_1 = \varrho_1 c^2$. The jump conditions may be simplified if an equation of state is known for the postshock gas,

$$p_2 = \kappa_2 (e_2 - \varrho_2 c^2), \quad (3.95)$$

where $\kappa_2 = 2/3$ for a nonrelativistic shock, and $\kappa_2 = 1/3$ for an ultrarelativistic shock. The shock strength may be described by relative velocity $\beta_{12} = (\beta_1 - \beta_2)/(1 - \beta_1 \beta_2)$ or $W_{12} = (1 - \beta_{12}^2)^{-1/2}$. The postshock gas satisfies $e_2/\varrho_2 c^2 = W_{12}$, and a convenient approximation for κ_2 is

$$\kappa_2 = \frac{1}{3} \left(1 + \frac{1}{W_{12}} \right). \quad (3.96)$$

It is exact if the gas is monoenergetic, i.e. particles have equal energies in the gas frame. Its error for Maxwellian gas is within 5%. Using equation (3.96), you can express all quantities in terms of β_{12} (or W_{12}), which remains as the only free parameter of the shock,

$$\beta_2 = \frac{\beta_{12}}{3}, \quad \varrho_2 = 4W_{12}\varrho_1, \quad e_2 = 4W_{12}^2 \varrho_1 c^2, \quad p_2 = \frac{4}{3} (W_{12}^2 - 1) \varrho_1 c^2. \quad (3.97)$$

These equations describe shocks of arbitrary strength, relativistic or nonrelativistic. The blast wave has two shocks, forward and reverse. The above equations with $W_{12} = \Gamma$ describe FS. The RS is described by the same equations when index 1 is replaced by 4 (unshocked ejecta) and index 2 is replaced by 3 (shocked ejecta). Pressures $p_f = p_2$ and $p_r = p_3$ are given by

$$p_f = \frac{4}{3} (\Gamma^2 - 1) \varrho_1 c^2, \quad p_r = \frac{4}{3} (W_{43}^2 - 1) \varrho_4 c^2. \quad (3.98)$$

Assume that the Lorentz factor of unshocked ejecta, Γ_{ej} , is known in the lab frame. If the pressure balance $p_f = p_r$ is assumed, it immediately determines the instantaneous Γ ,

$$\Gamma = \Gamma_{\text{ej}} \left[1 + 2\Gamma_{\text{ej}} \left(\frac{\varrho_1}{\varrho_{\text{ej}}} \right)^{1/2} \right]^{-1/2}. \quad (3.99)$$

This expression gives a quick estimate of Γ at the initial stage of the explosion, however, it is incorrect for the later evolution. In particular, energy conservation is not satisfied and requires a different solution.

Use the equations of motion for a spherical shell to derive the variation of the integrated mass-density Σ , enthalpy H and pressure P between RS and FS.

3.4. Relativistic Riemann Problem: Consider the one-dimensional special relativistic flow of a perfect fluid in the absence of a gravitational field. The Riemann problem then consists of computing the breakup of a discontinuity, which initially separates two arbitrary constant states L (left) and R (right) in the fluid. For classical hydrodynamics the solution can be found, e.g. in [396]. In the case of SRHD, the Riemann problem was considered by Marti and Müller, see [271], who derived an exact solution for the case of pure normal flow, generalizing previous results for zero initial velocities.

4 Relativistic Stellar Structure

In general relativity, a model of an isolated star (or other fluid ball, like boson stars) generally consists of a fluid-filled interior region, which is technically speaking a perfect fluid solution of the Einstein field equation, and an exterior region, which is an asymptotically flat vacuum solution. These two pieces must be carefully matched across the world sheet of a spherical surface, the surface of zero pressure. The exterior region of nonrotating compact objects is given in terms of the Schwarzschild solution. In this chapter we derive the famous Tolman–Oppenheimer–Volkoff (TOV) equations which give the hydrostatic equilibrium for relativistic stars.

4.1 Spacetime of Relativistic Stars

As a basic example of the above concepts of GR, we consider in this section the spacetime of nonrotating objects (Earth, Sun, stars, or black holes). Due to the high symmetries of these objects, all nondiagonal elements in the metric vanish, and, due to the static requirements for the gravitational fields, the metric elements are mere functions of the position of a spherically symmetric shell. Static and spherically symmetric nonrotating stars therefore generate a spacetime of the following form (for details, see [2, 18])

$$ds^2 = -\exp(2\Phi(r)) dt^2 + \exp(2\lambda(r)) dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) . \quad (4.1)$$

The coordinate r is a measure for the surfaces of the two-spheres, $4\pi r^2$, the coordinates θ and ϕ have the usual meaning of spherical coordinates, and the two functions $\Phi(r)$ and $\lambda(r)$ are uniquely given by the mass–energy distribution $\varrho(r)$ in the star. As in the Newtonian stellar structure, we can define the **total mass** inside the radius r

$$M(r) = 4\pi \int_0^r \varrho(r') r'^2 dr' . \quad (4.2)$$

The only difference is that this mass is given not only in terms of the mass-density, but in terms of the total mass–energy density ϱ which includes the internal energy density ϵ

$$\varrho \equiv \varrho_0 (1 + \epsilon/\varrho_0 c^2) . \quad (4.3)$$

The other difference is that the volume integration is not done with the proper volume element, $dV = 4\pi r^2 \exp \lambda(r) dr$, but with a somewhat reduced volume element. We can therefore define a second mass that is called **gravitational mass** M_G of the object

$$M_G = 4\pi \int_0^R \exp(\lambda(r)) \varrho(r') r'^2 dr'. \quad (4.4)$$

In particular, the **rest mass** M_0 of the object has to be defined with the proper volume element

$$M_0 = 4\pi \int_0^R \exp(\lambda(r)) \varrho_0(r) r^2 dr. \quad (4.5)$$

The entire structure of a compact star is then determined by the four equations called **Tolman–Oppenheimer–Volkoff (TOV) equations** (Oppenheimer and Volkoff [317])

$$\frac{dM(r)}{dr} = 4\pi \varrho(r) r^2 \quad (4.6)$$

$$\begin{aligned} \frac{dP(r)}{dr} = & -\frac{GM(r)\varrho(r)}{r^2} \left(1 + \frac{P(r)}{\varrho(r)c^2}\right) \\ & \times \left(1 + \frac{4\pi r^3 P(r)}{M(r)c^2}\right) \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1} \end{aligned} \quad (4.7)$$

$$e^{-2\lambda(r)} = 1 - \frac{2GM(r)}{c^2 r} \quad (4.8)$$

$$\frac{d\Phi(r)}{dr} = \frac{1}{1 - 2GM(r)/c^2 r} \left(\frac{GM(r)}{c^2 r^2} + \frac{4\pi GrP}{c^4} \right). \quad (4.9)$$

As in the Newtonian case, the total mass $M(r)$ inside a spherical shell of radius r also determines the hydrostatic equilibrium, but four corrections occur

- the mass-density ϱ_0 has to be replaced in terms of the total mass–energy density ϱ ;
- the inertial mass-density is given by $\varrho c^2 + P$ (see also equations of motion); this is the first correction factor on the right-hand side;
- pressure is an active volume correction (second factor);
- the metric of three-space enters in terms of the last factor; this factor is of particular importance, since it determines the stability properties of the solutions. The surface of the object always has to be far outside the Schwarzschild surface.

It is important to note that the curvature of three-space is entirely given in terms of the total mass, while the gravitational potential satisfies its Newtonian analogue, except for the inertial factor $\varrho c^2 + P$. It is then obvious that these structure equations go over into the Newtonian analog for $P \ll \varrho c^2$, i.e. roughly speaking for sound velocities much less than the velocity of light, for low compactness $2GM(r)/c^2 \ll r$ and for low pressure-mass $4\pi r^3 P(r) \ll M(r)c^2$. The compactness parameter has a particular

influence on the hydrostatic equilibrium (the last factor in the TOV equation). In this limit, three-space is flat, i.e. $\exp(\lambda) \equiv 1$ for all radii, and $\Phi(r) \equiv 1 + \Phi(r)/c^2$ with the following structure equations, usually derived in the theory of stellar structure,

$$\frac{dM(r)}{dr} = 4\pi\varrho_0(r) r^2 \quad (4.10)$$

$$\frac{dP(r)}{dr} = -\frac{GM(r)\varrho_0(r)}{r^2} \quad (4.11)$$

$$\frac{d\Phi(r)}{dr} = \frac{GM(r)}{r^2}. \quad (4.12)$$

Exterior Solution

As a further consequence we see that for vanishing pressure, i.e. in the **exterior region of a star** with $r > R$, the solution of the TOV equations is given by the **Schwarzschild solution** with

$$\exp(-2\lambda(r)) = 1 - \frac{2GM}{c^2r} \quad (4.13)$$

$$\exp(2\Phi(r)) = 1 - \frac{2GM}{c^2r}. \quad (4.14)$$

This Schwarzschild mass is now $M = 4\pi \int_0^R \varrho(r) r^2 dr$. The quantity $R_S = 2GM/c^2$ is called the Schwarzschild radius and scales with the mass

$$R_S = 2.98 \text{ km } (M/M_\odot). \quad (4.15)$$

4.2 Derivation of the TOV Equations

The TOV equations can be derived in a standard manner by calculating the Christoffel symbols in Schwarzschild coordinates and from there the Riemann and Ricci tensor. For highly symmetric spacetimes, there is, however, a method which is more suitable and physical than the coordinate based method. All physical quantities are expressed in terms of observer frames, also called tetrads (vierbein). Since this technique is the basic one, when we are dealing with rotating objects, we use it also here to calculate the Ricci tensor.

4.2.1 The Curvature of Static Spacetimes

In order to calculate the affine metric connection of this spacetime, we work with the local orthonormal one-form basis, which is naturally attached to the metric element

$$\begin{aligned} \Theta^0 &= \exp \Phi(r) dt, & \Theta^1 &= \exp \lambda(r) dr, \\ \Theta^2 &= r d\theta, & \Theta^3 &= r \sin \theta d\phi. \end{aligned} \quad (4.16)$$

This basis is orthonormal, so that the connection forms satisfy

$$\omega_{ab} + \omega_{ba} = 0, \quad \omega_{ab} = \eta_{ac}\omega_b^c, \quad a, b = 0, 1, 2, 3. \quad (4.17)$$

We first consider the external derivatives

$$d\Theta^0 = \Phi' \exp \Phi dr \wedge dt \quad (4.18)$$

$$d\Theta^1 = 0 \quad (4.19)$$

$$d\Theta^2 = dr \wedge d\theta \quad (4.20)$$

$$d\Theta^3 = \sin \theta dr \wedge d\phi + r \cos \theta d\theta \wedge d\phi. \quad (4.21)$$

Now we express the right-hand sides in terms of the orthonormal frames and obtain

$$d\Theta^0 = \Phi' \exp(-\lambda) \Theta^1 \wedge \Theta^0 \quad (4.22)$$

$$d\Theta^1 = 0 \quad (4.23)$$

$$d\Theta^2 = r^{-1} \exp(-\lambda) \Theta^1 \wedge \Theta^2 \quad (4.24)$$

$$d\Theta^3 = r^{-1} [\exp(-\lambda) \Theta^1 \wedge \Theta^3 + \cot \theta \Theta^2 \wedge \Theta^3]. \quad (4.25)$$

We compare this with the first Cartan structure equation (see Sect. 2.5.4)

$$d\Theta^a = -\omega_b^a \wedge \Theta^b \quad (4.26)$$

and can read off therefore the following connection coefficients

$$\omega_1^0 = \omega_0^1 = \Phi' \exp(-\lambda) \Theta^0 \quad (4.27)$$

$$\omega_2^0 = \omega_0^2 = \omega_3^0 = \omega_0^3 = 0 \quad (4.28)$$

$$\omega_1^2 = -\omega_2^1 = r^{-1} \exp(-\lambda) \Theta^2 \quad (4.29)$$

$$\omega_1^3 = -\omega_3^1 = r^{-1} \exp(-\lambda) \Theta^3 \quad (4.30)$$

$$\omega_2^3 = -\omega_3^2 = r^{-1} \cot \theta \Theta^3. \quad (4.31)$$

The second Cartan structure equation determines the curvature two-form Ω_a^b , which is also an element of the Lie algebra of the Lorentz group. The calculations are straightforward, as example we consider the forms

$$\Omega_1^0 = d\omega_1^0 + \omega_k^0 \wedge \omega_1^k = d\omega_1^0 \quad (4.32)$$

$$= (\Phi' \exp(-\lambda))' dr \wedge \Theta^0 + \Phi' \exp(-\lambda) d\Theta^0 \quad (4.33)$$

$$= (\Phi' \exp(-\lambda))' \exp(-\lambda) \Theta^1 \wedge \Theta^0 + (\Phi' \exp(-\lambda))^2 \Theta^1 \wedge \Theta^0 \quad (4.34)$$

$$\Omega_1^0 = -\exp(-2\lambda) [(\Phi')^2 - \Phi' \lambda' + \Phi''] \Theta^0 \wedge \Theta^1 \quad (4.35)$$

$$\Omega_2^0 = d\omega_2^0 + \omega_k^0 \wedge \omega_2^k = \omega_1^0 \wedge \omega_2^1 = -\exp(-2\lambda) \frac{\Phi'}{r} \Theta^0 \wedge \Theta^2 \quad (4.36)$$

In total, all six elements of the curvature two-form are given by

$$\Omega_1^0 = -\exp(-2\lambda)[(\Phi')^2 - \Phi'\lambda' + \Phi'']\Theta^0 \wedge \Theta^1 \quad (4.37)$$

$$\Omega_2^0 = -\frac{\Phi' \exp(-2\lambda)}{r} \Theta^0 \wedge \Theta^2 \quad (4.38)$$

$$\Omega_3^0 = -\frac{\Phi' \exp(-2\lambda)}{r} \Theta^0 \wedge \Theta^3 \quad (4.39)$$

$$\Omega_2^1 = \frac{\lambda' \exp(-2\lambda)}{r} \Theta^1 \wedge \Theta^2 \quad (4.40)$$

$$\Omega_3^1 = \frac{\lambda' \exp(-2\lambda)}{r} \Theta^1 \wedge \Theta^3 \quad (4.41)$$

$$\Omega_3^2 = \frac{1 - \exp(-2\lambda)}{r^2} \Theta^2 \wedge \Theta^3. \quad (4.42)$$

From this we can read off the Riemann tensor R^a_{bcd} , which is antisymmetric in both pairs of indices

$$\Omega_b^a = \frac{1}{2} R^a_{bcd} \Theta^c \wedge \Theta^d \quad (4.43)$$

and therefore the Einstein tensor follows from the Ricci tensor $R_{bc} = R^a_{bac}$

$$G_{ab} = R_{ab} - \frac{1}{2}\eta_{ab}R. \quad (4.44)$$

This yields the components

$$G_0^0 = \frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} - \frac{2\lambda'}{r} \right) = \frac{1}{r^2} \frac{d}{dr} [r(1 - \exp(-2\lambda))] \quad (4.45)$$

$$G_1^1 = \frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} + \frac{2\Phi'}{r} \right) \quad (4.46)$$

$$G_2^2 = G_3^3 = -\exp(-2\lambda) \left((\Phi')^2 - \Phi'\lambda' + \Phi'' + \frac{\Phi' - \lambda'}{r} \right) \quad (4.47)$$

and all other $G_{ab} = 0$. This is also a consequence of the high symmetry. The equality of G_2^2 and G_3^3 is a consequence of the isotropy on the sphere, but the radial component G_1^1 will in general differ from these. As a consequence, Einstein's equations provide three equations for the two functions $\Phi(r)$ and $\lambda(r)$. In fact, the third equation contains the hydrostatic equilibrium, since the equations of motion are not independent in general relativity.

4.2.2 Matter in the Interior

The matter in the interior of the star is described in terms of the energy-momentum tensor T^{ab} that assumes the form of a perfect fluid

$$T^{\mu\nu} = (\varrho c^2 + P) u^\mu u^\nu + P g^{\mu\nu}, \quad (4.48)$$

where $g^{\mu\nu}$ are the covariant components of the metric tensor. In the above equation, u^μ is the local fluid four-velocity $u^\mu = dx^\mu/d\tau$, where $d\tau = ds/c$. It satisfies the normalization $u^\mu u_\mu = -1$ (time-like vector field). ϱ is the total mass-energy density, and P the corresponding pressure. Because the star is static, the three-velocity of the vector field vanishes, and due to the normalization

$$u^0 = 1/\sqrt{-g_{00}} = 1/\alpha(r), \quad (4.49)$$

or $u_0 = -\alpha(r)$ is a measure of the redshift factor. For static stars the source of the gravitational field has the form

$$T_b^a = \text{diag} \{ -\varrho c^2, P, P, P \}. \quad (4.50)$$

This demonstrates that Einstein's field equations

$$G_b^a = \frac{8\pi G}{c^4} T_b^a \quad (4.51)$$

can explicitly be written as

$$G_0^0 = \frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} - \frac{2\lambda'}{r} \right) = \frac{8\pi G}{c^2} \varrho(r) \quad (4.52)$$

$$G_1^1 = \frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} + \frac{2\Phi'}{r} \right) = -\frac{8\pi G}{c^4} P(r) \quad (4.53)$$

$$\begin{aligned} G_2^2 = G_3^3 &= -\exp(-2\lambda) \left((\Phi')^2 - \Phi'\lambda' + \Phi'' + \frac{\Phi' - \lambda'}{r} \right) \\ &= -\frac{8\pi G}{c^4} P(r), \end{aligned} \quad (4.54)$$

where the first two equations provide us two independent equations for the functions $\Phi(r)$ and $\lambda(r)$

$$\frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} - \frac{2\lambda'}{r} \right) = \frac{8\pi G}{c^2} \varrho \quad (4.55)$$

$$\frac{1}{r^2} - \exp(-2\lambda) \left(\frac{1}{r^2} + \frac{2\Phi'}{r} \right) = -\frac{8\pi G}{c^4} P. \quad (4.56)$$

The first equation is equivalent to

$$(r \exp(-2\lambda))' = 1 - \frac{8\pi G}{c^2} \varrho r^2, \quad (4.57)$$

which can be integrated with the asymptotic flatness condition to yield the three-space metric

$$\exp(-2\lambda) = 1 - \frac{2GM(r)}{c^2r}$$

(4.58)

with the total mass inside radius r given as

$$M(r) \equiv 4\pi \int_0^r \varrho(r') r'^2 dr'. \quad (4.59)$$

By subtracting the second equation from the first one, we obtain

$$\exp(-2\lambda) (\Phi' + \lambda') = \frac{4\pi G}{c^2} (\varrho + P) r. \quad (4.60)$$

This is equivalent to

$$\Phi' = \frac{1}{1 - 2GM(r)/c^2r} \left(\frac{GM(r)}{c^2r^2} + \frac{4\pi G}{c^4} rP \right). \quad (4.61)$$

This demonstrates now, how the gravitational force is generalized in GR. In particular, pressure is a source of the gravitational field, and the Schwarzschild metric acts as a modification in the denominator of the force law.

We now solve the first equation for λ' and set $c = 1 = G$

$$-2r\lambda' = (1 - 8\pi r^2 \varrho) \exp(2\lambda) - 1 \quad (4.62)$$

and also solve the second equation for Φ'

$$2r\Phi' = (1 + 8\pi r^2 P) \exp(2\lambda) - 1 \quad (4.63)$$

and take the derivative of this last equation and multiply by r

$$2r\Phi' + 2r^2\Phi'' = [2r\lambda' (1 + 8\pi r^2 P) + (16\pi r^2 P + 8\pi r^3 P')] \exp(2\lambda) \quad (4.64)$$

We solve this equation for Φ'' using once again the above two relations

$$2r^2\Phi'' = 1 + (16\pi r^2 P + 8\pi r^3 P') \exp(2\lambda) - (1 + 8\pi r^2 P) (1 - 8\pi r^2 \varrho) \exp(4\lambda). \quad (4.65)$$

We can also square the above relations for Φ' to obtain

$$2r^2(\Phi')^2 = \frac{1}{2} (1 + 8\pi r^2 P)^2 \exp(4\lambda) - (1 + 8\pi r^2 P) \exp(2\lambda) + \frac{1}{2}. \quad (4.66)$$

These equations provide relations for Φ' , λ' , Φ'' and Φ'^2 in terms of ϱ , P , P' and $\exp(2\lambda)$, which can be expressed entirely in terms of the included mass $M(r)$. Therefore, all the metric functions can be eliminated from the third of Einstein's equations by substitution of the above results. In this way, we obtain after some lengthy calculations the equation for the **relativistic hydrostatic equilibrium**

$$\boxed{\frac{dP}{dr} = -\frac{GM(r)\varrho(r)}{r^2} \left(1 + \frac{P(r)}{\varrho(r)c^2}\right) \times \left(1 + \frac{4\pi r^3 P(r)}{M(r)c^2}\right) \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1}.} \quad (4.67)$$

This equation could also be obtained directly from the equations of motion $T_{r;r}^r = 0$. This equation is, however, a mere consequence of Einstein's equation and should not be considered as an independent equation. This demonstrates once again that $\Phi(r)$ is the analogue of the Newtonian potential $\Phi(r)$

$$P' = -(\varrho c^2 + P) \Phi'. \quad (4.68)$$

A further consequence of this derivation is a relation for the gravitational force

$$\frac{d\Phi}{dr} = \frac{GM(r)}{c^2 r^2} \frac{1 + 4\pi r^3 P/M(r)c^2}{1 - 2GM(r)/c^2 r}. \quad (4.69)$$

For a given equation of state $P = P(\varrho)$, the TOV equations can easily be integrated from the origin with initial conditions $M(0) = 0$ and an arbitrary value for the central density $\varrho_c = \varrho(0)$, until the pressure $P(r)$ will vanish at some radius R . **To each possible equation of state, there is a unique family of stars parametrized by the central density, i.e. we obtain a sequence of stellar models $M = M(\varrho_c)$.**

4.2.3 The Exterior Schwarzschild Solution

For radii exceeding the mass distribution, $r \geq R_*$, where $r = R_*$ with $P(R_*) = 0$ denotes the surface of the star and $M = M(R_*)$ the mass of the star, the solution is simply given by

$$\exp(-2\lambda) = 1 - \frac{2GM}{c^2 r} \quad (4.70)$$

$$\Phi' = \frac{GM}{r^2 (1 - 2GM/c^2 r)}. \quad (4.71)$$

The second equation can be integrated with the boundary condition $\exp[\Phi(r)] \rightarrow 1$ for $r \rightarrow \infty$

$$\boxed{\exp[2\Phi(r)] = 1 - \frac{2GM}{c^2 r}, \quad r \geq R_*}. \quad (4.72)$$

It can be shown that these solutions are compatible with the third Einstein equation $G_2^2 = 0$.

This is the famous Schwarzschild solution with its metric

$$\boxed{ds^2 = -\left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2 d\Omega^2,} \quad (4.73)$$

uniquely determined by the mass M of the central object. The quantity

$$R_S = 2GM/c^2 = 3 \text{ km } M/M_\odot \quad (4.74)$$

is called the **Schwarzschild radius**. R_*/R_S is therefore a measure for the compactness of the star. When $r \rightarrow R_S$, the metric becomes degenerate, and redshift from this surface becomes infinite. Objects are no longer stable, when their surface is too near to the Schwarzschild surface – they will ultimately collapse. For this reason, neutron stars have a maximally possible mass, $M_c \simeq 2 M_\odot$. Objects with higher masses will collapse for ever and form a **black hole**, where the surface with $r = R_S$ appears as a horizon for observers at infinity (i.e. a surface of infinite redshift).

The Schwarzschild metric will become singular at the radius $r = R_S$, but, as shown by the expressions for the Riemann tensor, this tensor stays finite there, since, for example, $\Phi' \exp(-2\lambda) = GM/r^2$. This has the consequence that the Riemann tensor diverges for $r \rightarrow 0$ as, e.g. $R_{202}^0 = -GM/r^3$. This is now a point singularity which contains all the mass measured by the parameter M . Physically, it is clear that this is a mathematical singularity which will be removed, hopefully, by some quantum gravity theory.

4.2.4 Stable Branches for Degenerate Stars

Not all branches of a sequence $M = M(\varrho_c)$ are stable. This can be tested by means of radial oscillations. Degenerate stars with $dM/d\varrho_c < 0$ are found to be unstable and will finally collapse towards Neutron stars, or black holes [15]. The corresponding Sturm–Liouville problem for the oscillation modes has been worked out for the first time by Chandrasekhar in 1964 [115]. As a consequence, we find by solving the TOV equations for a given EoS a maximum possible mass with a corresponding central maximal density.

For degenerate EoS two sequences are found – on the one hand the **white dwarfs** where the electron pressure is equilibrating gravity, and at higher densities the **neutron stars** where the degenerate pressure of neutrons and quarks is responsible for the equilibrium. A third family of pure quark stars (or strange stars) might exist, but the observations are not yet conclusive.

4.2.5 Metric for Relativistic Stars

The structure of relativistic stars has attracted much attention since the formulation of general relativity. Theoretical models of relativistic stars have first been considered by Tolman [394], and Oppenheimer and Volkoff [317].

The solutions of the TOV equations are numerically found by integrating outward from the origin, $r = 0$, towards the surface where pressure vanishes. They are parameterized by the central density ϱ_c , and the adiabatic constant Γ in the EoS, or by an explicit numerical EoS. A given solution of the TOV equations, or a relativistic star, is characterized by its mass M , and its radius R . The profiles of the two

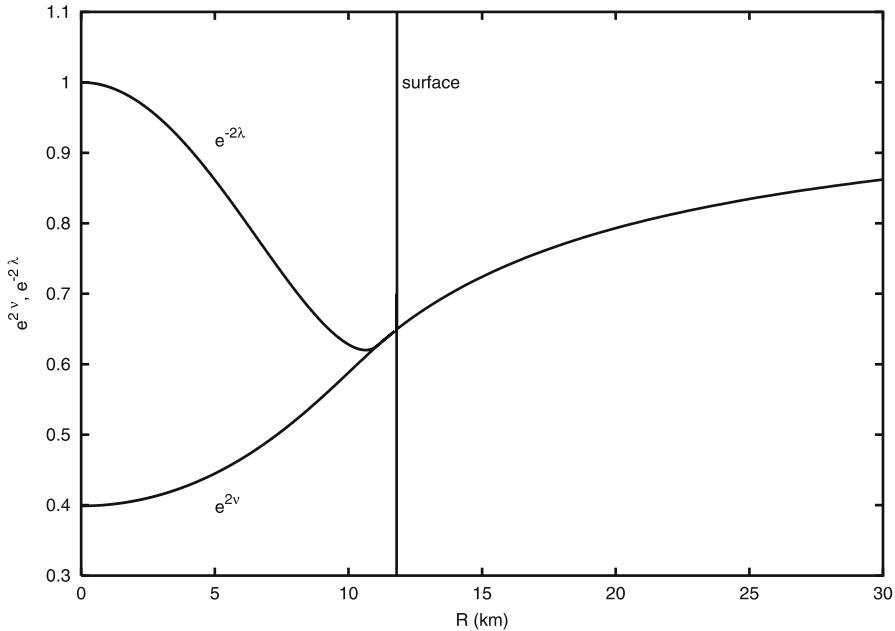


Fig. 4.1. Metric functions of static neutron stars as a function of radius in units of km. The redshift factor $\alpha = \exp \nu = \exp \Phi$ steadily increases from the center of the star towards the asymptotic region. The metric function $\exp[-2\lambda(r)] = 1 - 2GM(r)/c^2$ is flat near the center and reaches a minimum near the surface of the star, where it joins the redshift factor. Figure provided by A. Bauswein (ZAH, Landessternwarte)

metric functions are shown in Fig. 4.1 for a typical neutron star. The redshift factor $\alpha(r) = \exp \nu(r)$ steadily increases from the center of the star towards the asymptotic region. The metric function $\exp[-2\lambda(r)] = 1 - 2GM(r)/c^2$ is flat near the center and reaches a minimum near the surface of the star, where it joins the redshift factor.

4.3 A Variational Principle for the Stellar Structure

It is remarkable that the TOV equations for the stellar structure can be derived from a variational principle (for details, see Weinberg [20]):

A particular stellar configuration, with uniform entropy per nucleon and chemical composition, will satisfy the TOV equations for equilibrium, if and only if the quantity M defined by

$$M = \int_0^\infty 4\pi\varrho(r) r^2 dr \quad (4.75)$$

is stationary with respect to all variations of $\varrho(r)$ that leave unchanged the total number of baryons, $n(r) \equiv \varrho_0(r)/m_u$ as the baryon density,

$$N_B = \int_0^\infty 4\pi r^2 n(r) \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1/2} dr. \quad (4.76)$$

To prove this theorem one uses the Lagrange multiplier method (we follow here the elegant outline given by Weinberg [20]): M will be stationary with respect to all variations that leave N_B fixed provided there exists a constant λ for which $M - \lambda N_B$ is stationary with respect to all variations. In general, we get

$$\begin{aligned} \delta M - \lambda \delta N_B &= \int_0^\infty 4\pi r^2 \delta \varrho(r) dr \\ &\quad - \lambda \int_0^\infty 4\pi r^2 dr \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1/2} \delta n(r) \\ &\quad - \lambda \frac{G}{c^2} \int_0^\infty 4\pi r \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-3/2} n(r) \delta M(r) dr. \end{aligned} \quad (4.77)$$

These variations are supposed not to change the entropy per nucleon, i.e.

$$0 = \delta \left(\frac{\varrho}{n}\right) + P \delta \left(\frac{1}{n}\right) \quad (4.78)$$

or

$$\delta n(r) = \frac{n(r)}{\varrho(r) + P(r)} \delta \varrho(r). \quad (4.79)$$

And in addition we have

$$\delta M(r) = \int_0^r 4\pi r'^2 \delta \varrho(r') dr'. \quad (4.80)$$

If we interchange the r and r' integration in the last term of the total variation

$$\begin{aligned} \delta M - \lambda \delta N_B &= \int_0^\infty 4\pi r^2 \left[1 - \frac{\lambda n(r)}{\varrho + P} \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1/2}\right] \\ &\quad - \lambda \frac{G}{c^2} \int_r^\infty 4\pi r' n(r') \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-3/2} dr' \delta \varrho(r) dr. \end{aligned} \quad (4.81)$$

Thus $\delta M - \lambda \delta N_B$ will vanish if and only if

$$\begin{aligned} \frac{1}{\lambda} &= \frac{n(r)}{\varrho + P} \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1/2} \\ &\quad + \frac{G}{c^2} \int_r^\infty 4\pi r' n(r') \left(1 - \frac{2GM(r)}{c^2 r}\right)^{-3/2} dr'. \end{aligned} \quad (4.82)$$

This will be the case for some multiplier λ if and only if the right-hand side is independent of r , i.e. only and only if

$$0 = \left[\frac{n'}{\varrho + P} - \frac{n(P' + \varrho')}{(\varrho + P)^2} \right] \left(1 - \frac{2GM(r)}{c^2r} \right)^{-1/2} + \frac{Gn/c^2}{\varrho + P} \left(4\pi\varrho - \frac{GM(r)}{c^2r^2} \right) \left(1 - \frac{2GM(r)}{c^2r} \right)^{-3/2} - 4\pi Grn \left(1 - \frac{2GM(r)}{c^2r} \right)^{-3/2}. \quad (4.83)$$

The condition for uniform entropy gives

$$\frac{d}{dr} \left(\frac{\varrho}{n} \right) + P \frac{d}{dr} \left(\frac{1}{n} \right) = 0, \quad (4.84)$$

and therefore

$$n'(r) = \frac{n(r)\varrho'(r)}{\varrho(r) + P(r)}. \quad (4.85)$$

Therefore δM vanishes for all $\delta\varrho(r)$ that give $\delta N_B = 0$ provided

$$-r^2 P' = \frac{G}{c^2} (\varrho + P) (M(r) + 4\pi r^3 P) \left(1 - \frac{2GM(r)}{c^2r} \right)^{-1/2} \quad (4.86)$$

and this is the TOV equation.

Problems

4.1. Perihelion Advance in Schwarzschild: Use the energy equation (8.71) to derive the following relation for $u = L^2/GMr$

$$\left(\frac{du}{d\phi} \right)^2 + \frac{L^2}{G^2 M^2} - 2u + u^2 - \frac{2G^2 M^2}{L^2} u^3 = \frac{E^2 L^2}{G^2 M^2}. \quad (4.87)$$

A circular orbit is then located at $u = 1$. By differentiation with respect to u we obtain a second-order equation

$$\frac{d^2u}{d\phi^2} - 1 + u = \frac{3G^2 M^2}{L^2} u^2. \quad (4.88)$$

Expand u into a Newtonian solution plus a small perturbation

$$u = u_N + u_1 \quad (4.89)$$

with the Newtonian solution given by

$$\frac{d^2 u_N}{d\phi^2} - 1 + u_N = 0, \quad (4.90)$$

and a first-order deviation by

$$\frac{d^2 u_1}{d\phi^2} + u_1 = \frac{3G^2 M^2}{L^2} u_N^2. \quad (4.91)$$

The Newtonian solution is the classical expression

$$u_N = 1 + e \cos \phi, \quad (4.92)$$

where e is given by

$$e^2 = 1 - \frac{L^2}{G M a}. \quad (4.93)$$

Derive from the perturbed equation the perihelion advance per orbit

$$\Delta\phi = \frac{6\pi G^2 M^2}{L^2} = \frac{6\pi G M}{c^2 (1 - e^2) a}.$$

(4.94)

Calculate this value for Mercury per orbit and transform it to its value per century. Calculate the periastron shift for a star S_3 orbiting the central black hole in the Galactic center [358].

4.2. Motions in Schwarzschild: Derive the Christoffel symbols for the Schwarzschild geometry and the equations of motion for test particles in terms of the affine parameter τ .

4.3. 3+1 Split of Schwarzschild: Schwarzschild spacetime has a vanishing extrinsic curvature. Use the Poisson equation (2.380) and the Ricci equations (2.388) to derive the TOV equations.

4.4. Incompressible Stars: Solve the TOV equations for an incompressible star (constant density ϱ_* in the interior). Show that the hydrostatic equilibrium leads to the solution for the pressure

$$P(r) = \varrho_* \left[\frac{R \sqrt{R - 2GM} - \sqrt{R^3 - 2GMr^2}}{\sqrt{R^3 - 2GMr^2} - 3R\sqrt{R - 2GM}} \right], \quad (4.95)$$

and the solution for the redshift factor

$$\alpha(r) = \frac{3}{2} \sqrt{1 - 2GM/R} - \frac{1}{2} \sqrt{1 - 2GMr^2/R^3} \quad (4.96)$$

for $r < R$, where R is the radius of the star. Show that static solutions are only possible for a compactness $GM/R < 4/9$.

4.5. TOV Equation: Write a simple code to integrate the TOV equation in the case of a polytropic equation of state $P = K\varrho^\Gamma$, for given central pressure P_c and polytropic index Γ .

5 White Dwarfs

A white dwarf is what stars like our Sun become when they have exhausted their nuclear fuel. Near the end of its nuclear burning stage, such a star expels most of its outer material (creating a planetary nebula), until only the hot core remains, which then settles down to become a very hot ($T > 100,000$ K) young white dwarf. Since a white dwarf has no way to keep itself hot unless it is accreting matter from a nearby star (as a cataclysmic variable), it cools down over the course of the next billion years. Many nearby, young white dwarfs have been detected as sources of soft X-rays (i.e. lower-energy X-rays); recently, soft X-ray and extreme ultraviolet observations have become a powerful tool in the study of the composition and structure of the thin atmosphere of these stars.

A typical white dwarf is half as massive as the Sun, yet only slightly bigger than the Earth. This makes white dwarfs one of the densest forms of matter, surpassed only by neutron stars. Once a star is degenerate, gravity cannot compress it any more because quantum mechanics tells us there is no more available space to be taken up. A white dwarf survives therefore, not by internal combustion, but by quantum mechanical principles that prevent its complete collapse. Such degenerate matter has other unusual properties; for example, the more massive a white dwarf is, the smaller it is, contrary to what is observed for normal stars. This is because the more mass a white dwarf has, the more its electrons must squeeze together to maintain enough outward pressure to support the extra mass. There is a limit on the amount of mass a white dwarf can have, however. It was found by Subrahmanyan Chandrasekhar to be 1.4 times the mass of our Sun, and is called the *Chandrasekhar limit* after its discoverer.

With a surface gravity of 100,000 times that of the Earth, the atmosphere of a white dwarf is very strange. The heavier atoms in its atmosphere sink and the lighter ones remain at the surface. Some white dwarfs have almost pure hydrogen or helium atmospheres, the lightest of elements. Also, the very strong gravity pulls the atmosphere close around it in a very thin layer. Underneath the atmosphere of many white dwarfs, we think there is a 50 km thick crust, the bottom of which is a crystalline lattice of carbon and oxygen atoms. One might make the comparison between a cool carbon/oxygen white dwarf and a diamond.

Since white dwarf stars glow just from residual heat, the oldest white dwarfs will be the coldest and thus the faintest. By searching for faint white dwarfs, one can estimate the length of time the oldest white dwarfs have been cooling. The

luminosity function of white dwarfs at low luminosities, and especially the position of its cutoff, provides important information about the age of the Galactic disk.

5.1 Observations of Isolated White Dwarfs

5.1.1 Sirius B

There are several ways to observe white dwarf stars. The first white dwarf ever to be discovered was found because it is a companion star to Sirius, a bright star near the constellation Canis Major. In 1844, astronomer Friedrich Bessel noticed that Sirius had a slight back and forth motion, as if it were being orbited by an unseen object. In 1863, this mysterious object was finally resolved by optician Alvan Clark and it was found to be a white dwarf. This pair is now referred to as Sirius A and B, B being the white dwarf. The orbital period of this system is about 50 years. Since white dwarfs are very small and thus very hard to detect, binary systems are a helpful way to locate them. As with the Sirius system, if a star seems to have some sort of unexplained motion, we may find that the single star is really a multiple system. Upon close inspection we may find that it has a white dwarf companion.

The black-body spectrum of Sirius B peaks at 110 nm, corresponding to a temperature of 27,000 K (Fig. 5.3). From the known absolute magnitude (the distance of the system is 8.6 lightyears), the radius is calculated as 4200 km, smaller than the Earth, but as massive as the Sun.

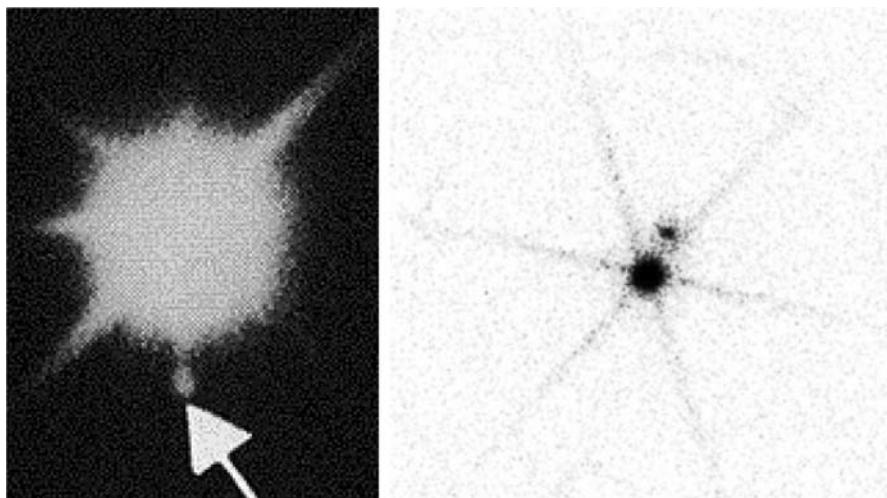


Fig. 5.1. The binary system Sirius A and B. *Left:* optical image (arrow points to the white dwarf); *right:* Chandra image (here the white dwarf is the bright object). Image courtesy: Chandra Observatory

Chandra Resolves Sirius B

The Chandra image shows two sources and a spike-like pattern due to the support structure for the transmission grating (Fig. 5.1). The bright source is Sirius B, a white dwarf star that has a surface temperature of about 27,000 degrees Kelvin, which produces very low energy X-rays. The dim source at the position of Sirius A, a normal star more than twice as massive as the Sun, may be due to ultraviolet radiation from Sirius A leaking through the filter on the detector.

5.1.2 Field White Dwarfs and Classification

Optical spectra of white dwarfs have been classified according to their dominant element in the atmosphere

- DA: strong hydrogen lines
- DB: strong He I lines
- DO: strong He II lines
- DC: no strong lines (continuous) spectrum
- DZ: strong metal lines (excluding carbon)
- DQ: strong carbon lines.

Multiple families are shown in decreasing order, e.g. DAB, DQAB.

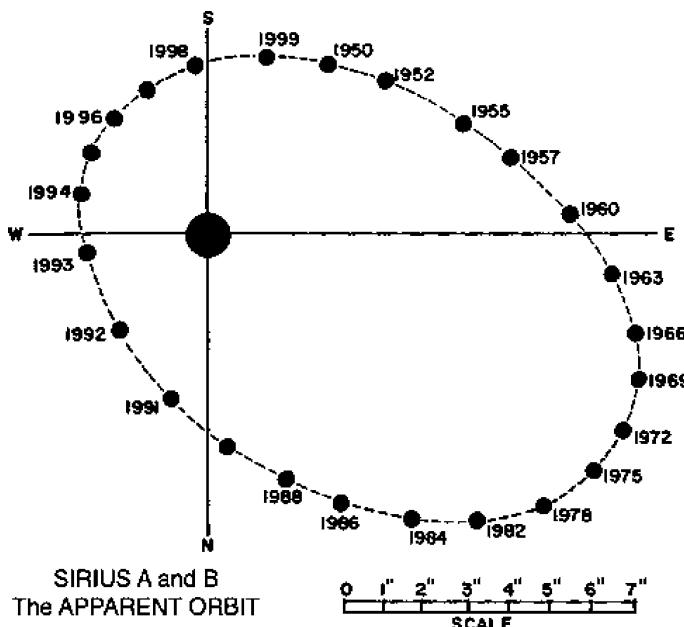


Fig. 5.2. The orbit of the binary system Sirius A and B. Binary parameters: $a_A = 6.43$ AU, $a_B = 13.4$ AU, $e = 0.592$, $i = 136.5$ deg, $M_A = 2.14 M_{\odot}$, $M_B = 1.03 M_{\odot}$, $P_b = 50.1$ yrs

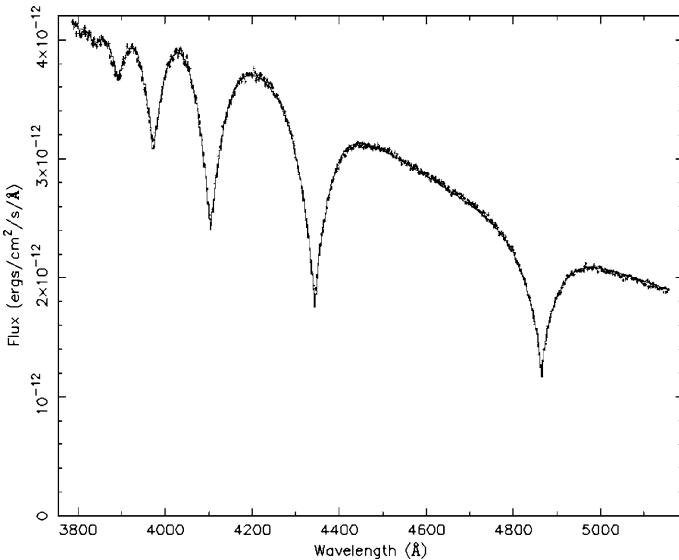


Fig. 5.3. HST Balmer line spectrum of Sirius B with the best-fit synthetic spectrum (*solid line*) corresponding to $T_{\text{eff}} = 25,193$ K and $\log g = 8.556$. Figure adapted from Barstow et al. [59]

Exact distances have been determined for about 20 white dwarfs with the Hipparcos satellite (Table 5.1).

Surface Compositions

The surface composition is quite well known from spectroscopic observations:

- 80% of all WDs are DAs.
- Most WDs have pure or nearly pure H or He atmospheres.
- DAs are found from hottest to coolest WDs.
- Non-DAs start with hot stars:
 - DOs are found for $T_{\text{eff}} > 45,000$ K with He II or He I;
 - DBs for $T_{\text{eff}} < 30,000$ K, He I only;
 - DCs (featureless) for $T_{\text{eff}} < 11,000$ K;
 - No He-rich WDs between 45,000 and 30,000 K.

Spectroscopic Features

The strong gravity of white dwarfs results in rapid settling of heavier elements, e.g. hydrogen always rises to the top and can mask other elements. Given a white dwarf atmosphere modelling is generally considered to be more tractable than for other stars. If trace elements are seen, as in DZ white dwarfs, then they must be of recent origin (e.g. accretion from the ISM, comets, etc.).

Table 5.1. Parallax data for white dwarfs measured by Hipparcos and by ground-based observations. Practically all WDs are of the DA-type

WD number	HIP name	HIP	Spectral type	Hipparcos parallax [mas]	Ground-based parallax [mas]	V magnitude
0046+051	vMa 2	3829	DZ	226.95 ± 5.35	232.5 ± 1.9	12.371 ± 0.018
0148+467	GD 279	8709	DA	63.08 ± 3.79	61.0 ± 7.0	12.440 ± 0.030
0227+050	Feige 22	11650	DA	41.15 ± 4.96	45.0 ± 5.0	12.799 ± 0.0014
0232+035	Feige 24	12031	DA	13.44 ± 3.62	13.1 ± 2.5	12.411 ± 0.003
0310-688	LB 3303	14754	DA	98.50 ± 1.46	84.9 ± 15.0	11.387 ± 0.019
0426+588	Stein 2051B	21088	DC	181.36 ± 3.67	180.6 ± 0.8	12.440 ± 0.030
0501+527	D 191-B2B	23692	DA	14.53 ± 3.09	23.3 ± 2.2	11.781 ± 0.0055
0644+375	He 3	32560	DA	64.91 ± 3.37	66.2 ± 2.1	12.057 ± 0.006
0713+584	GD 294	35307	?	-1.89 ± 2.97		11.980 ± 0.030
1134+300	GD 140	56602	DA	65.28 ± 3.61	70.4 ± 10.9	12.487 ± 0.019
1142-645	L 145-141	57367	DQ	216.40 ± 2.11	218.3 ± 6.7	11.503 ± 0.017
1314+293	HZ 43	64766	DA	31.26 ± 8.33	15.5 ± 3.4	12.914 ± 0.030
1327-083	Wolf 485	65877	DA	55.50 ± 3.77	61.8 ± 2.8	12.313 ± 0.005
1337+705	G 238-44	66578	DA	40.33 ± 2.89	30.5 ± 5.9	12.792 ± 0.004
1544-377	L 481-60	77358	DA	65.60 ± 0.77	73.5 ± 9.4	12.800 ± 0.030
1620-391	CD-3810980	80300	DA	78.04 ± 2.40	65.5 ± 7.6	11.010 ± 0.011
1647+591	G226-29	82257	DA	91.13 ± 2.33	81.9 ± 4.0	12.240 ± 0.031
1917-077	LDS 678A	95071	DB	89.08 ± 7.16	99.2 ± 2.5	12.280 ± 0.030
2032+248	Wolf 1346	101516	DA	67.65 ± 2.32	69.4 ± 2.3	11.528 ± 0.001
2039-202	L 711-10	102207	DA	47.39 ± 4.04	42.4 ± 8.4	12.330 ± 0.020
2149+021	G 93-48	107968	DA	39.84 ± 4.47	40.8 ± 2.5	12.738 ± 0.008

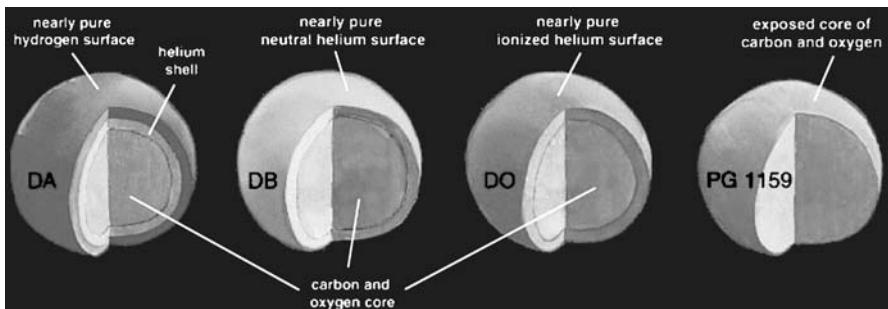


Fig. 5.4. Classification of white dwarfs

White Dwarf Search

All galaxy surveys will also provide a set of WDs. Kleinman et al. [229] have published a catalog of spectroscopically identified WDs in the first data release of the Sloan Digital Sky Survey. The Sloan Digital Sky Survey (SDSS) is a continuing imaging and spectroscopic survey of some 7–10 thousand square degrees in the

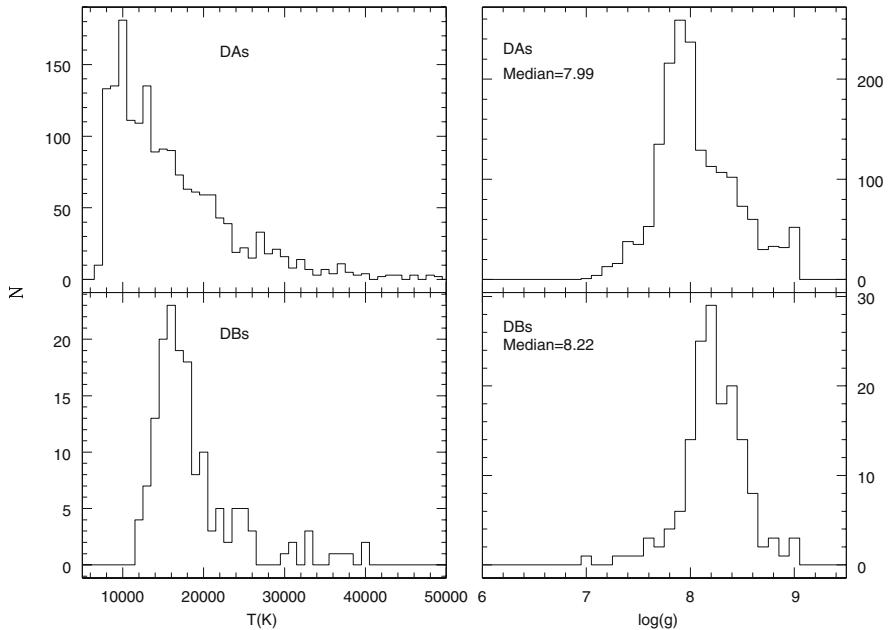


Fig. 5.5. White dwarf stars found in the SDSS with spectra fitted to atmospheric models. *Left:* distribution in effective temperature for DAs (top) and DBs (bottom); *right:* distribution in $\log g$ for the same types. Figure adapted from Kleinman et al. [229]

north Galactic cap. Though its main focus is extragalactic, there are many Galactic spin-off projects resulting from the survey. In an area which is 1400 square degrees, they find 2551 white dwarf stars of various types and an additional 144 objects as uncertain white dwarf stars. Of all white dwarf stars, 1888 are nonmagnetic DA types and 171 nonmagnetic DBs. The remaining 492 objects consist of all different types: DO, DQ, DC, DZ and hybrid stars. The DA and DB spectra are fitted with a grid of atmospheric models to determine T_{eff} and $\log g$ for each object (Fig. 5.5). This catalog nearly doubles the known sample of spectroscopically identified white dwarf stars.

Mass Distribution

The typical mass of field white dwarfs is $\simeq 0.6 M_{\odot}$ (Fig. 5.6). Masses M and radii R of the 1833 WDs in the SDSS sample were directly determined from the effective temperature and $\log g$ [265]. Given T_{eff} and $\log g$ of a star, mass and radius are computed from the definition of the surface gravity

$$M = \frac{R^2}{G} 10^{\log g} \quad (5.1)$$

and from the mass–radius relation $M = M(R, T_{\text{eff}})$. The resulting mass distribution (Fig. 5.6) exhibits both the peak of $0.562 M_{\odot}$ and the high mass tail.

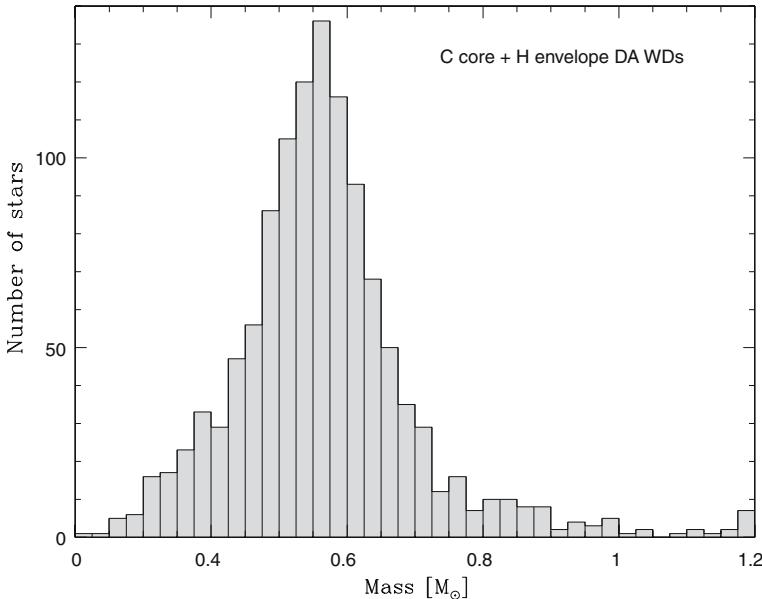


Fig. 5.6. Histogram of the mass distribution from the SDSS sample of Kleinman [265]

The peak mass of the SDSS catalog is in excellent agreement with earlier studies, which used much less numerous samples and various methods of white dwarf mass determination (for details, see [265]).

5.1.3 White Dwarfs in Globular Clusters

Some very nearby white dwarf stars can be observed directly through telescopes, though they are extremely faint. M4, shown in Fig. 5.7, is the nearest globular cluster to the Earth. It contains hundreds of thousands of stars visible with ground-based telescopes, and is expected to contain about 40,000 white dwarfs. This globular cluster formed early in the history of the Milky Way, and today is a veritable stellar retirement community. It is so ancient (about 13 billion years old) that all of its stars that began with 80% or more of the Sun's mass have already evolved off the main sequence to become red giants, and many have turned into white dwarfs. In the Hubble Space Telescope picture shown in Fig. 5.7, the brightest of the detected white dwarfs (the faint pin-pricks circled on the right-hand side of the figure) is no more luminous than a 100-watt light bulb seen at the Moon's distance.

5.1.4 Magnetic White Dwarfs

There are over 65 catalogued isolated magnetic white dwarfs, comprising about 2% of the total WD population (see Wickramasinghe and Ferrario [418], Fig. 5.8). Their

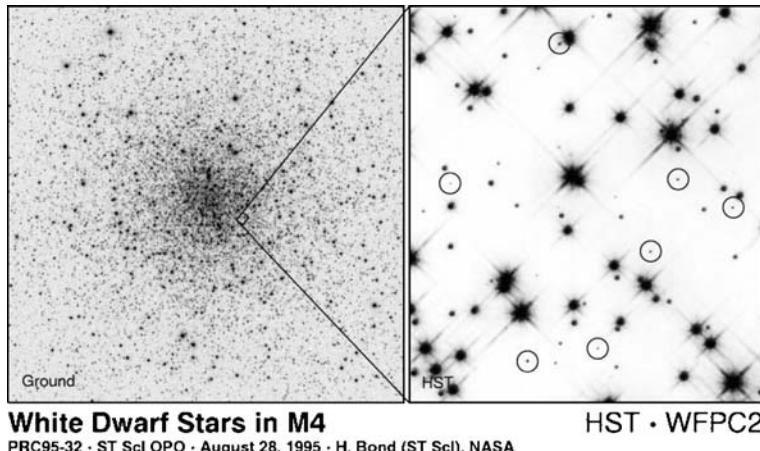


Fig. 5.7. White dwarfs in the globular cluster M4. The globular cluster M4 is the nearest globular cluster to the Earth. It contains hundreds of thousands of stars visible with ground-based telescopes, and is expected to contain about 40,000 white dwarfs [HST image archive]

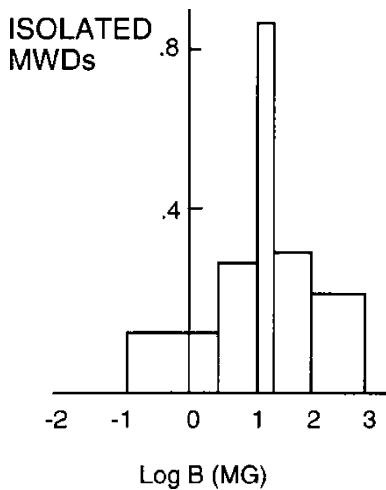


Fig. 5.8. Normalized histogram of magnetic white dwarfs

field strengths range from 30 kG to 1000 MG with temperatures in the range of 4000 K to 50,000 K. These magnetic WDs are extremely useful for measuring spin periods. A significant fraction of magnetic WDs display photometric, spectroscopic or polarimetric variability.

Remember that the mean field strength on the surface of the Sun is only a few gauss, but at the bottom of the convective zone magnetic fields of the order of 100 kG are expected. Since magnetic flux is conserved during the contraction phase, in principle strong fields could occur on white dwarf stars. On average, these white dwarfs have larger mass, and some rotate rapidly, but others not at all. Magnetism

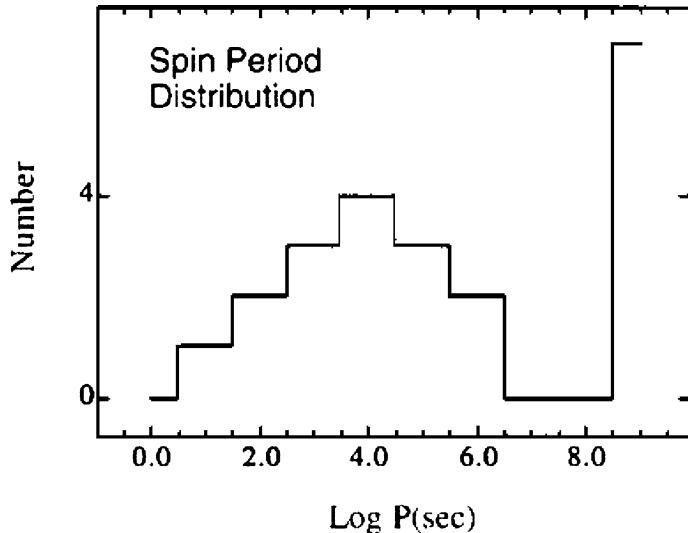


Fig. 5.9. Period distribution of magnetic white dwarfs. All stars with periods greater than 10 years are in the last bin

thus influences in a way the time evolution of spinning white dwarfs. The very origin of these strong fields is, however, still mysterious. A dynamo process would require rapid rotation – and this is in general not observed.

Traditional measurements of rotation rates for stars – photometric variations from spots and rotational broadening of otherwise narrow absorption lines – fail for white dwarfs. **White dwarfs are spotless stars.** Photometric time series do not show evidence of rotational modulation by spots. With very high gravities producing lines that are 10 nm wide, natural broadening mechanisms swamp rotational broadening – even for velocities close to break-up. The discovery of sharp cores in the spectral lines of some DA WDs and of nonradial pulsations have produced a handful of useful WD rotation velocities. Magnetic WDs with magnetic fields in excess of 100 kG, show evidence for time-variability and periodicities of the magnetic features, providing in this way a measurement of their rotation velocities (Fig. 5.9). For more on these fascinating objects, see Wickramasinghe and Ferrario [418].

5.1.5 Ultracool White Dwarfs as Cosmochronometers

White dwarf cooling theory (see Sect. 5.6) presents many fascinating aspects from a physical point of view, but, for the present purposes, it is sufficient to study Fig. 5.10 which illustrates the evolutionary tracks of five representative models of white dwarfs in the HR diagram. Except for their different mass, these models are similar in that they all consist of a pure C core, surrounded by a He mantle containing 10^{-2} of the total mass of the star, and an outermost H layer containing 10^{-4} of the

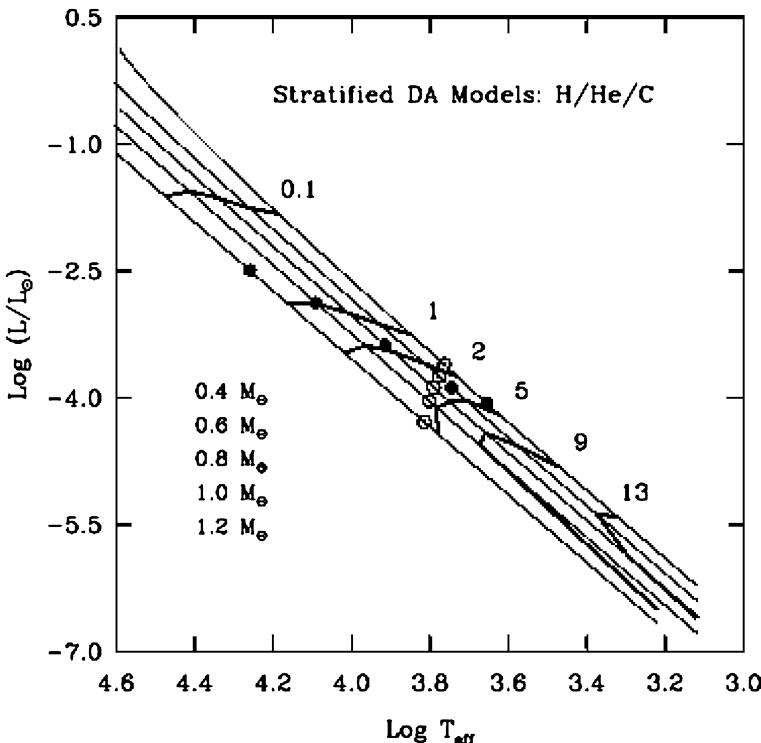


Fig. 5.10. White dwarf tracks in the Hertzsprung–Russell diagram for different masses. The numbers indicate the age in Gyr

total mass of the star. These values roughly correspond to the maximum amounts of H and He that can survive the previous, hot planetary nebula phase.

The detailed evolutionary tracks nearly follow curves of constant radii (straight lines with a negative slope in this log–log version of the HR diagram), particularly at low luminosities. Furthermore, these tracks indicate that the more massive white dwarfs are also those that are the smaller. As we will see, this peculiar mass–radius relation is a consequence of electron degeneracy supporting the white dwarfs against gravitational collapse.

While the evolutionary paths of cooling white dwarfs in the HR diagram are extremely simple, it should be noticed that the cooling time to a given luminosity say, is not only a function of T_{eff} , but also a strong function of the total mass of the star. This is shown in Fig. 5.10 by the heavy curves representing isochrones. At relatively high luminosities, the plot shows that a more massive white dwarf takes longer to cool to a given T_{eff} than a less massive object. The reason is that the more massive star has a larger energy reservoir: there are more C ions with energy kT . However, there is a dramatic reversal of behavior in the cooler phases of the evolution. Because of their larger masses and smaller radii, more massive white dwarfs have larger internal densities (for comparable temperatures) and, therefore, develop a crystallized core

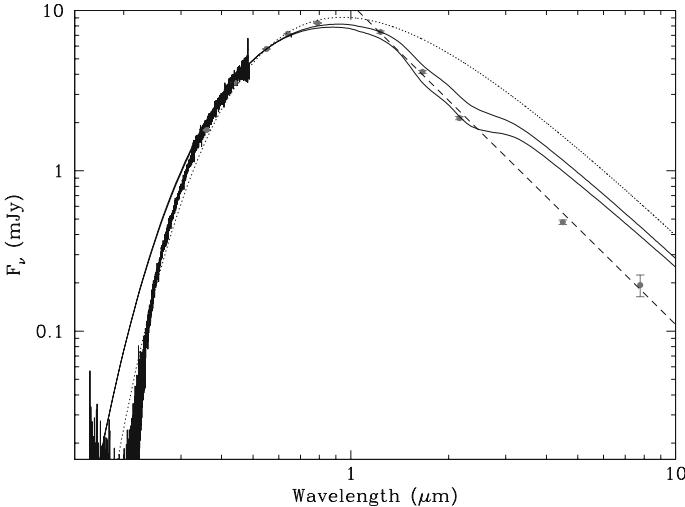


Fig. 5.11. Spectral energy distribution of LHS 1126, along with a 5400 K black-body (dotted line). White dwarf models with $T_{\text{eff}} = 5400$ K, $\log g = 7.9$, and $\log N(\text{He})/N(\text{H}) = 1.5$ (lower solid line) and $\log N(\text{He})/N(\text{H}) = 1$ (upper solid line) are also shown. The dashed line shows a power law with $\alpha = 2$. Dots are data from the Spitzer Telescope. Figure adapted from Kilic et al. [226]

earlier, at higher effective temperatures. This is explicitly illustrated in Fig. 5.10, where the small filled circle on each track indicates the onset of crystallization at the center of each evolving model. Note that crystallization is a first-order phase transition and, consequently, is accompanied by the release of latent heat. This extra source of energy produces a delay in the cooling of a crystallizing white dwarf.

The Sloan Digital Sky Survey (SDSS) has increased the number of known field cool white dwarfs from tens of objects to thousands. The Spitzer Space Telescope opened a new window into the Universe by enabling accurate mid-infrared photometry of faint objects (microjansky-level sensitivity). In order to understand the CIA opacity, and other unrecognized sources of opacity in cool white dwarf atmospheres, the Spitzer Space Telescope has been used to observe nearby, relatively bright, cool white dwarfs [226], where mid-infrared photometry for 18 cool white dwarfs including LHS 1126 has been presented. Cool white dwarfs have atmospheres dominated by hydrogen or helium. Both hydrogen and helium are neutral below 5000 K and the primary opacity source in H-rich cool ($T_{\text{eff}} \leq 5500$ K) white dwarf atmospheres is believed to be collision induced absorption (CIA) of molecular hydrogen. H-rich white dwarfs are predicted to become redder as they cool until the effects of CIA become significant below 5500 K.

Spitzer observations demonstrate that all H-rich white dwarfs with $T_{\text{eff}} < 7000$ K show slight mid-infrared flux deficits. Having several stars with small deficits makes these deficits significant. Moreover, LHS 1126 shows significantly depressed mid-infrared fluxes relative to white dwarf models (Fig. 5.11).

WDs in Solar Neighborhood

The usefulness of white dwarfs as cosmochronometers had already been firmly established in 1987 when Winget et al. [421] first demonstrated that the white dwarf population in the solar neighborhood – a population characteristic of the galactic disk – could be used to estimate independently the age of the disk. Over the years, the method has been refined through improvements in the quality of the observational material available and improvements in the cooling models.

Figure 5.12 illustrates how the age of the local disk can be estimated through a comparison of the observed and theoretical luminosity functions of local white dwarfs. On the observational side, Leggett, Ruiz, and Bergeron (1998) and Knox, Hawkins, and Hambly (1999) published their studies of the luminosity function of white dwarfs in the solar neighborhood. Compared to older samples, these samples have provided much improved estimates of effective temperatures, bolometric corrections, and absolute magnitudes. It contains 43 objects and constitutes a complete proper motion survey. In comparison, the Knox et al. (1999) is also a complete survey, but it is a colorimetric survey. It contains 58 objects.

Recently, a sample of white dwarfs has been selected from SDSS DR3 imaging data using their reduced proper motions, based on improved proper motions from

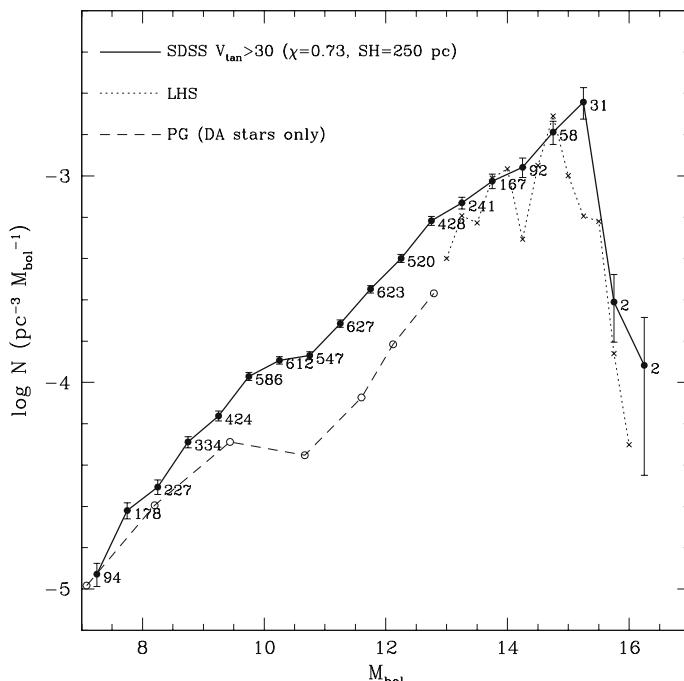


Fig. 5.12. The luminosity function of local white dwarfs derived from the SDSS survey. The numbers indicate the numbers of stars for each data point. The *dashed line* at the bright end is from Liebert et al. [255], based on the analysis from the PG Survey, including DA WDs only

SDSS plus USNO-B combined data [193]. Numerous SDSS and followup spectra are used to quantify completeness and contamination of the sample; kinematic models are used to understand and correct for velocity-dependent selection biases. A luminosity function is constructed covering the range $7 < M_{\text{bol}} < 16$ (Fig. 5.12). This white dwarf luminosity function now based on 6000 stars is remarkably smooth, and rises nearly monotonically to $M_{\text{bol}} = 15.3$. It then drops abruptly, although the small number of low-luminosity stars in the sample and their unknown atmospheric composition prevent quantitative conclusions about this decline. These surveys clearly suggest the existence of a **bump in the luminosity function**, peaking around 10^{-4} solar units in luminosity. Interestingly, this bump, or excess of white dwarfs, is naturally expected from theory and corresponds to the delays in cooling associated with the release of latent heat upon crystallization and the effects of convective coupling.

The simplest explanation for the observed drop-off of the density of white dwarfs at low luminosities, and the one that has been accepted quite generally, is that the first white dwarfs that were formed in the disk and that are now in our neighborhood, are still bright enough to be visible. Most of them, with representative or average masses have piled up at a luminosity $\simeq 10^{-4} L_{\odot}$, while the more massive of them, much less numerous, have trickled down through Debye cooling to lower luminosities during the same time and populate the tail at the faint end of the luminosity function.

A comparison of the curves in Fig. 5.12 with the observed points, particularly, the coolest bin, suggests an age of 11 Gyr or less for the local disk. It is important to realize here that this estimate is related to the assumption of a pure C core composition in the cooling models used in this illustrative example. A core composition containing a mixture of C and O, as is actually expected from stellar evolution theory, would lead to a smaller value than obtained here for the age of the disk. This is because the specific heat of a gram of oxygen is less than the specific heat of a gram of carbon under the fluid/solid physical conditions encountered in white dwarf interiors.

GAIA is an ambitious space mission, adopted within the scientific programme of the European Space Agency (ESA) in October 2000. Its main purpose is to measure the positions and proper motions of an extremely large number of objects with unprecedented accuracy. GAIA will have a great impact on our understanding of the Galactic white dwarf population. The superb astrometric capabilities of GAIA will provide us with an unprecedented number of white dwarfs with excellent astrometric measurements. In particular, the disk white dwarf population will be probed up to distances of 400 pc, with typical errors smaller than 10%, both in proper motion and parallax and with a completeness ranging from nearly 100% for objects within 100 pc to 30% for objects within 400 pc [397]. Thus, GAIA will determine with high accuracy the disk white dwarf luminosity function and its drop-off. This excellent situation will, however, not pertain for the halo white dwarf population. GAIA will also provide very precise information on the physical mechanisms (crystallization and phase separation) important during the cooling process by comparing the theoretical luminosity functions of disk white dwarfs

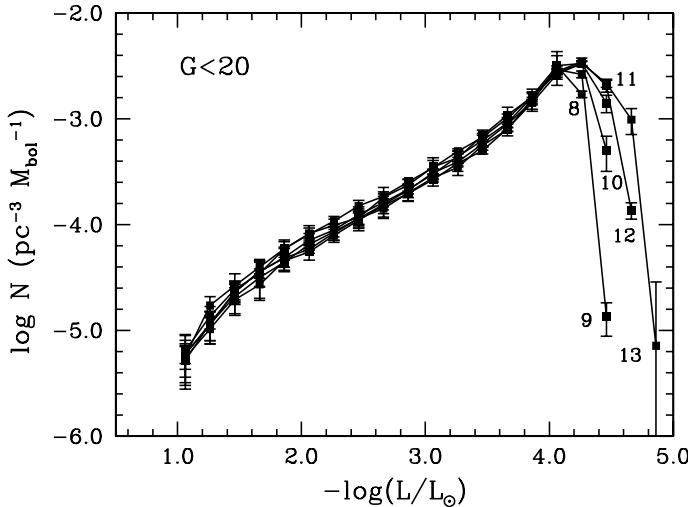


Fig. 5.13. Luminosity function of disk white dwarfs for several ages of the disk expected in GAIA observations, ranging from 8 to 13 Gyr, with an interval of 1 Gyr. The *error bars* are the standard deviation of the 40 independent Monte Carlo realizations. Figure adapted from [397]

with the observations (see Fig. 5.13). In addition, the luminosity function of massive disk white dwarfs will constrain the star formation history of the Galactic disk.

White Dwarf Populations in Distant Clusters

To study faint white dwarf populations in distant systems such as open and globular clusters, one often has to deal with what could be called “minimal” or two-band photometry that produces a single color–magnitude diagram (CMD). Cooling theory can be used in conjunction with model atmospheres to compute the evolutionary tracks and plot isochrones in the CMD. However, the photometric scatter in such CMD’s is generally large for white dwarfs and only qualitative results can be obtained from the direct comparison of isochrones with observational points in the CMD’s. Clearly, the proper way to exploit the information contained in these diagrams is through actual stellar counts and the construction of observed luminosity functions.

The observational signature of the finite age of the white dwarf population in a cluster is the maximum in white dwarf density, the expected pile-up at some luminosity characteristic of the cluster, followed by a drop-off in number density at lower luminosities. Obviously, the sensitivity of the observations must be large enough to reveal this pile-up; otherwise, white dwarf cosmochronology can be used only to provide lower limits to the age of a cluster. Recently, the luminosity function for WDs in M4 has been derived from HST observations [192].

5.2 What is Inside a White Dwarf?

To say that white dwarfs are strange is an understatement. An Earth-sized white dwarf has a density of $1 \times 10^9 \text{ kg/m}^3$. In comparison, the Earth itself has an average density of only $5.4 \times 10^3 \text{ kg/m}^3$. That means a white dwarf is a million times as dense.

Because a white dwarf is no longer able to create internal pressure, gravity unopposedly crushes it down until even the very electrons that make up a white dwarf's atoms are mashed together. Under normal circumstances, identical electrons (those with the same spin) are not allowed to occupy the same energy level. Since there are only two ways an electron can spin, only two electrons can occupy a single energy level. This is what is known in physics as the Pauli exclusion principle. And in a normal gas, this is not a problem; there are not enough electrons floating around to completely fill up all the energy levels. But in a white dwarf, all of its electrons are forced close together; soon all the energy levels in its atoms are filled up with electrons. If all the energy levels are filled, and it is impossible to put more than two electrons in each level, than our white dwarf has become degenerate. For gravity to compress the white dwarf, it must force electrons where they cannot go. Once a star is degenerate, gravity cannot compress it any more because quantum mechanics tells us there is no more available space to be taken up. Our white dwarf therefore survives, not by internal combustion, but by quantum mechanical principles that prevent its complete collapse.

Degenerate matter has other unusual properties; for example, the more massive a white dwarf is, the smaller it is. This is because the more mass a white dwarf has, the more its electrons must squeeze together to maintain enough outward pressure to support the extra mass. There is a limit on the amount of mass a white dwarf can have, however. It was found by Subrahmanyan Chandrasekhar to be 1.4 times the mass of our Sun, and is is called the **Chandrasekhar limit** after its discoverer.

With a surface gravity of 100,000 times that of the Earth, $g = GM/R^2 \simeq 10^8$ in cgs units, the atmosphere of a white dwarf is very strange. The heavier atoms in its atmosphere sink and the lighter ones remain at the surface. Some white dwarfs have almost pure hydrogen or helium atmospheres, the lightest of elements. Also, the very strong gravity pulls the atmosphere close around it in a very thin layer, that, if were it on Earth, would be lower than the tops of our skyscrapers. Underneath the atmosphere, there is a 50 km thick **crust**, the bottom of which is a crystalline lattice of carbon and oxygen atoms. One might make the comparison between a cool carbon/oxygen white dwarf and a diamond (Fig. 5.14).

While spectroscopic measurements of white dwarfs have made significant advances in our understanding of the composition and structure of white dwarf atmospheres, interpreting some of these results relies on theoretical models of the evolutionary paths the stars follow and, in particular, the theoretical **mass-radius relation**. This was first defined by Chandrasekhar in his Nobel prize winning work on white dwarf structure. However, he and later authors only considered a fully degenerate configuration. Recent work has realized that the nondegenerate envelope plays an important role, particularly in the hot white dwarfs which form the EUV

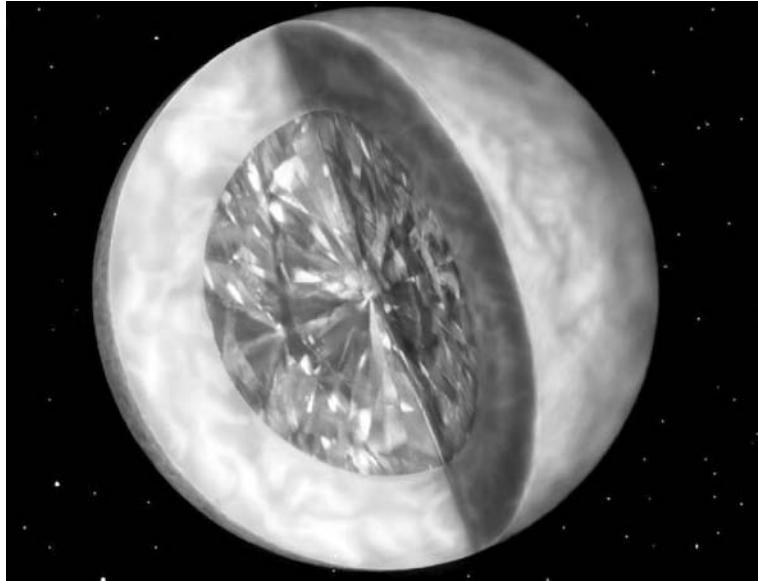


Fig. 5.14. Internal structure of old C/O white dwarfs. The core of a cool white dwarf consists of a C/O crystallized lattice (a kind of gigantic diamond), surrounded by a thick crust consisting of He and H and a small H atmosphere. Image credit: Travis Metcalfe and Ruth Bazinet, CfA

and X-ray emitting samples. Whichever framework is used, the relationship between the mass and radius of a white dwarf remains a theoretical concept which has hardly been tested empirically due to the difficulty of making accurate measurements of mass and radius in such faint objects.

Onset of Degeneracy in Stellar Structure

The electron degeneracy pressure will be important, if the phase space available for the electrons becomes minimal. For this purpose we estimate the mean momentum difference between two electrons in a Maxwell gas

$$\Delta p_e = \sqrt{\langle (\mathbf{p}_1 - \mathbf{p}_2)^2 \rangle} = \sqrt{2\mathbf{p}_1^2} \simeq \sqrt{6m_e k_B \bar{T}} \simeq \sqrt{\frac{12m_e G M_* \mu m_H}{7R}} \quad (5.2)$$

for $\Gamma = 5/3$. Here, we used the gravitational potential energy W of a star

$$W = - \int_0^R \frac{GM(r)}{r} \varrho_0 4\pi r^2 dr = -3 \int_0^R P 4\pi r^2 dr. \quad (5.3)$$

Together with the equation of state of an ideal gas, $P = \varrho_0 k_B T / (\mu m_u)$, we obtain

$$W = -\frac{3M_*}{\mu m_u} k_B \bar{T}, \quad (5.4)$$

where \bar{T} denotes the mean stellar temperature (defined over the mass). On the other hand, the gravitational energy is always

$$W = -\frac{6}{7} \frac{GM_*^2}{R_*} \quad (5.5)$$

with the factor $6/7$ characteristic for polytropic states with $\Gamma = 5/3$. In addition, the mean interparticle distance is given by

$$\Delta q_e \simeq n_e^{-1/3} = \left(\frac{\mu_e m_u}{\varrho} \right)^{1/3} \simeq \left(\frac{4\mu_e m_u R_*^3}{M_*} \right)^{1/3}. \quad (5.6)$$

From this we can calculate the phase-space volume available for the electrons in a star of mass M_* with radius R_*

$$\begin{aligned} (\Delta q_e \Delta p_e)^3 &\simeq 4\mu_e \left(\frac{12\mu}{7} \right)^{3/2} \left(\sqrt{Gm_e R_*} m_u^{5/6} M_*^{1/6} \right)^3 \\ &\simeq 40 \left(1 \times 10^{-26} \left(\frac{M_*}{M_\odot} \right)^{1/6} \left(\frac{R_*}{R_\odot} \right)^{1/2} \text{g cm}^2 \text{s}^{-1} \right)^3 \\ &\simeq 180 h^3 \left(\frac{M_*}{M_\odot} \right)^{1/2} \left(\frac{R_*}{R_\odot} \right)^{3/2} > h^3. \end{aligned} \quad (5.7)$$

For a star with $M_* = M_\odot$, the phase space of electrons will exceed h^3 , provided its radius exceeds some critical value, $R_* \geq 3 \times 10^{-2} R_\odot$. For the case

$$(\Delta q_e \Delta p_e)^3 \simeq h^3 \quad (5.8)$$

the Pauli exclusion principle is important, and therefore degeneracy sets in. This estimate also shows that the degenerate pressure of electrons is important in Brown Dwarfs and Jupiter type planets. The equation of state for these low-density objects is quite complicated.

5.3 Equation of State below the Neutron Drip Density

The pressure and energies in white dwarfs and neutron stars are nonthermal – thermal effects due to a finite temperature can be treated as a perturbation. One may treat high-density matter as having zero temperature. The equation of state (EoS) then reduces to a single parameter function, $P(\varrho_0)$ and $\varrho(\varrho_0)$, where P is the pressure, ϱ_0 is the rest-mass density and $\varrho = \varrho_0(1 + \epsilon/\varrho_0 c^2)$ is the total mass–energy density which accounts for the internal (possibly) relativistic particle energies as well as the rest-mass energy.

There exist two main regimes of high density. As long as all nucleons are confined to nuclei, their contribution to the total pressure is negligible compared to that of the degenerate electrons. At some threshold density, $\varrho_{n\text{-drip}}$, it becomes favorable for

the nuclei to disintegrate, i.e. the neutrons drip out of the nuclei and form a nucleon gas. The standard EoS of Baym, Pethick and Sutherland (BPS [64]) suggests that $\rho_{\text{n-drip}} \simeq 4 \times 10^{11} \text{ g cm}^{-3}$. One therefore distinguishes between the EoS below neutron drip and above neutron drip density. For white dwarfs we merely need the EoS below the neutron drip.

In matter below the neutron drip the ions provide a Coulomb lattice of point-like charges, which is to a good approximation independent of the surrounding electrons. The EoS of this matter is then governed mainly by the electron gas, and we may treat these electrons in a first approximation as an ideal fermion gas, where at extremely low densities Coulomb corrections have to be included. For high densities corrections enter through the inverse beta-decay just below the neutron drip. The standard equations for cold, degenerate matter in white dwarfs (helium, carbon, oxygen, and possibly iron dominated models) have been derived by Chandrasekhar (1931) and Hamada and Salpeter [190]. Models for **equilibrated matter**¹ have been derived by Dirac (1930) and Feynman, Metropolis and Teller (1932) for low densities, $\rho \leq 10^4 \text{ g cm}^{-3}$.

The Ideal Fermion Gas

In this approximation, the electrostatic energy associated with the structure of matter is much smaller than the Fermi energies. Coulomb forces are therefore generally negligible to a first approximation. We treat the electron component by a cold single species of gas of noninteracting fermions. At zero temperature, the fermions fill all the states with momentum $p \leq p_F$ and none of the states with $p > p_F$, where p_F is the Fermi momentum of the particles. The corresponding Fermi energy is

$$E_F = \sqrt{(p_F c)^2 + (mc^2)^2}, \quad (5.9)$$

where m is the fermion mass and c the speed of light.

For electrons, the number density n_e is directly related to their Fermi momentum, $p_{F,e}$, by integrating over all occupied phase space

$$n_e = \int_0^{p_{F,e}} n_e(p) d^3 p = \frac{2}{h^3} \int_0^{p_{F,e}} 4\pi p^2 dp = \frac{8\pi p_{F,e}^3}{3h^3}. \quad (5.10)$$

$h = 6.63 \times 10^{-27}$ erg s is Planck's constant. The factor 2 arises from the spin degeneracy of the electrons. The occupation number $n_e(p)$ is a dimensionless Lorentz-invariant function.

The pressure the electrons supply is calculated through the mean momentum flux of the electron gas (here given for an isotropic phase-space distribution)

¹ The equilibrium isotope of matter is the nucleus of highest binding energy per nucleon; at low densities this isotope is normally $^{56}_{26}\text{Fe}$, but as the density increases, the atomic mass and the neutron to proton ratio also increase.

$$P = \frac{1}{3} \int_0^{p_{F,e}} v_e(p) p n_e(p) d^3 p = \frac{2}{h^3} \int_0^{p_{F,e}} \frac{p^2 c^2}{\sqrt{p^2 c^2 + (mc^2)^2}} 4\pi p^2 dp. \quad (5.11)$$

In this expression, $v_e(p) \equiv p_e c^2 / E_e$ is the velocity of the electron. For the following it is useful to introduce the **electron Compton wavelength** $\Lambda_e \equiv h/(2\pi m_e c)$, as well as the dimensionless Fermi momentum $x \equiv p_{F,e}/m_e c$. In terms of these parameters, the integral can be performed to yield

$$P = \frac{m_e c^2}{\Lambda_e^3} \Phi(x), \quad (5.12)$$

where

$$\Phi(x) = \frac{1}{8\pi^2} \left[x \sqrt{1+x^2} (2x^2/3 - 1) + \ln \left(x + \sqrt{1+x^2} \right) \right]. \quad (5.13)$$

The mass–energy density is also related to the Fermi momentum

$$\begin{aligned} \epsilon_e &= \int_0^{p_{F,e}} E_e(p) n_e(p) d^3 p \\ &= \frac{2}{h^3} \int_0^{p_{F,e}} \sqrt{p^2 c^2 + (mc^2)^2} 4\pi p^2 dp = \frac{m_e c^2}{\Lambda_e^3} \chi(x), \end{aligned} \quad (5.14)$$

where

$$\chi(x) = \frac{1}{8\pi^2} \left[x \sqrt{1+x^2} (1+2x^2) - \ln \left(x + \sqrt{1+x^2} \right) \right]. \quad (5.15)$$

While the degenerate electrons contribute most of the pressure and internal energy of white dwarf matter, the mass-energy is dominated by the ions, which are completely nonrelativistic at these densities. Thus, the density of matter can be expressed in terms of the electron number density

$$\varrho = \varrho_0 = \frac{m_B n_e}{Y_e} \quad (5.16)$$

where Y_e is the mean number of electrons per nucleon and m_B is the mean nucleon mass

$$m_B = \frac{\sum_i m_i n_i}{\sum_i A_i n_i}. \quad (5.17)$$

For carbon C, $m_B = 1.66057 \times 10^{-24}$ g (atomic mass unit m_u). In the case of white dwarfs, we have to good approximation $Y_e = Z/A = 0.5$ (Z : atomic number, A : atomic weight). This is appropriate for fully ionized helium, carbon or oxygen. Combining the three equations for ϱ_0 , P and ϵ_e provides the basic EoS of electron-pressure dominated high-density condensed matter. Thereby, the Fermi momentum

has to be expressed in terms of the density, $p_{F,e} \propto \varrho_0^{1/3}$, or in terms of the dimensionless quantity x , $Y_e = 0.5$,

$$\varrho_0 = 1.95 \times 10^6 \text{ g cm}^{-3}. \quad (5.18)$$

Sometimes, the **mean molecular weight per electron**, μ_e , is introduced

$$\mu_e = \frac{m_B}{m_u Y_e} \simeq 2, \quad (5.19)$$

so that

$$\boxed{\varrho_0 = \mu_e m_u n_e = 0.97395 \times 10^6 \mu_e x^3 \text{ g cm}^{-3}} \quad (5.20)$$

or

$$\boxed{x = 1.0088 \times 10^{-2} \left(\frac{\varrho_0}{\mu_e} \right)^{1/3}.} \quad (5.21)$$

The density is expressed in terms of cgs units.

These expressions show that the electrons become relativistic at densities of a million grams per cc, $x \geq 1$. At higher densities, the phase space of the electrons is squeezed that much that the electrons must move relativistically. It is therefore instructive to look at these limits for the EoS, i.e. for $x \ll 1$ and $x \gg 1$, respectively. It is easy to derive the corresponding asymptotic expansions, first, for $x \ll 1$

$$\Phi(x) = \frac{1}{15\pi^2} \left[x^5 - \frac{5}{14}x^7 + \frac{5}{24}x^9 + \dots \right] \quad (5.22)$$

$$\chi(x) = \frac{1}{3\pi^2} \left[x^3 + \frac{3}{10}x^5 - \frac{3}{56}x^7 + \dots \right] \quad (5.23)$$

and for the extreme relativistic limit, $x \gg 1$,

$$\Phi(x) = \frac{1}{12\pi^2} \left[x^4 - x^2 + \frac{3}{2} \ln(2x) + \dots \right] \quad (5.24)$$

$$\chi(x) = \frac{1}{4\pi^2} \left[x^4 + x^2 - \frac{1}{2} \ln(2x) + \dots \right]. \quad (5.25)$$

In these two limits, we can write the leading terms in the form of a **polytropic** equation of state

$$\boxed{P = K \varrho_0^\Gamma} \quad (5.26)$$

with the following values for the constants:

- in the nonrelativistic case, $\Gamma = 5/3$ and (in cgs units)

$$K = \frac{3^{2/3}\pi^{4/3}}{5} \frac{\hbar^3}{m_e m_u^{5/3} \mu_e^{5/3}} = 1.0036 \times 10^{13} Y_e^{5/3} \quad (5.27)$$

- in the extreme relativistic case, $\Gamma = 4/3$,

$$K = \frac{3^{1/3} \pi^{2/3}}{4} \frac{\hbar c}{m_u^{4/3} \mu_e^{4/3}} = 1.2435 \times 10^{15} Y_e^{4/3}. \quad (5.28)$$

Since the constant K is in cgs units, it provides the pressure in dyne cm⁻² for a density in g cm⁻³. The composition of matter only enters over Y_e . Helium, carbon and oxygen all have the same equation of state, while that of iron ($Y_e = 26/56 = 0.43$) is somewhat softer. Fully equilibrated matter (often called catalyzed) has a Y_e that decreases with density, the EoS is therefore softer than matter composed of a single entity.

This free electron pressure EoS is a good approximation below the neutron drip. It was employed by Chandrasekhar (1930), for which he received the Nobel prize in 1983. In later years more exact treatments were considered with two main corrections – electrostatic effects at low densities and neutronization (i.e. inverse beta-decay) at higher densities.

Electrostatic Corrections

The ideal EoS for fermions is somewhat modified by electrostatic corrections due to the fact that the local distribution of charge is very nonuniform. Positive charge is concentrated in ions which causes the average electron-ion separation to be smaller than the average distance between the electrons. The electric potential felt by the electrons is then attractive, and this reduces the pressure for a given density.

These electrostatic corrections to the cold equation of state are mostly important at relatively low densities. Electrostatic energies are inversely proportional to the average separation between the particles, $\langle r \rangle \propto n_e^{-1/3}$. The relative importance between a degenerate, nonrelativistic electron and an ion of charge Z can be estimated

$$\frac{E_C}{E_F} = \frac{Ze^2/\langle r \rangle}{p_{F,e}^2/2m_e} \propto n_e^{-1/3} \quad (5.29)$$

since $E_F = p_{F,e}^2 \propto n_e^{2/3}$ is the Fermi kinetic energy of the nonrelativistic electron. This shows that the relative importance of electrostatic corrections decrease with density as $n_e^{-1/3}$. Numerically, this ratio follows more exactly as

$$\frac{E_C}{E_F} = 2 \left(\frac{1}{3\pi^2} \right)^{2/3} \frac{Z}{a_0} \frac{1}{n_e^{1/3}} = \left(\frac{n_e}{Z^3 \times 6 \times 10^{22} \text{ cm}^{-3}} \right)^{-1/3}. \quad (5.30)$$

Here, $a_0 = \hbar^2/m_e c^2$ is the Bohr radius.

For low temperatures, the ions are located in a lattice that maximizes the inter-ion separation.

In the Wigner-Seitz approximation, this lattice is considered as built up by spherical cells of size $4\pi r_0^3/3 = 1/n_N$, where n_N is the number density of nuclei. In this approximation, the gas is imagined to be divided up into neutral spheres of

radius r_0 about each nucleus, which contains the Z electrons closest to the nucleus. This is a bad approximation in the lab, but is justified for the higher densities in white dwarfs.

One can estimate the total Coulomb energy of a cell

$$E_C = E_{e-e} + E_{e-i} = -\frac{9}{10} \frac{Z^2 e^2}{r_0}, \quad (5.31)$$

where E_{e-e} represents the energy to assemble a uniform sphere of Z electrons, and E_{e-i} the energy to assemble the electron sphere about the nucleus. Since these cells are neutral, interactions between electrons and nuclei of different cells are ignored. The electrostatic energy per electron is then

$$\frac{E_C}{Z} = -\frac{9}{10} \left(\frac{4\pi}{3} \right)^{1/3} Z^{2/3} e^2 n_e^{1/3}. \quad (5.32)$$

In this approximation, the number density of electrons is

$$n_e = \frac{Z}{4\pi r_0^3 / 3}. \quad (5.33)$$

From this we can calculate the pressure

$$P_C = n_e^2 \frac{d(E_C/Z)}{dn_e} = -\frac{3}{10} \left(\frac{4\pi}{3} \right)^{1/3} Z^{2/3} e^2 n_e^{4/3}. \quad (5.34)$$

In the nonrelativistic limit we have for the Chandrasekhar pressure

$$P_0 \rightarrow \hbar^2 (3\pi^2)^{2/3} \frac{n_e^{5/3}}{5m_e} \quad (5.35)$$

so that for the total pressure $P = P_0 + P_C$

$$\frac{P}{P_0} = 1 - \frac{Z^{2/3}}{2^{1/3} \pi a_0 n_e^{1/3}}. \quad (5.36)$$

This would predict $P = 0$ for

$$n_e^{\text{crit}} = \frac{Z^2}{2\pi^3 a_0^3}, \quad (5.37)$$

corresponding to a density, $A \simeq 2Z$,

$$\varrho_0^{\text{crit}} \simeq 0.4 Z^2 \text{ g cm}^{-3}. \quad (5.38)$$

This corresponds to $\varrho_0 \simeq 250 \text{ g cm}^{-3}$ for iron, instead of the lab value 7.86 g cm^{-3} . This discrepancy is due to the fact that this approximation breaks down for low

densities. In fact, an accurate equation of state at lab densities is very complicated to derive, because electron shell effects mask the simpler statistical effects. Above a few times lab densities, however, a statistical approach to the equation of state works quite well. This approach is sufficient to treat the low mass range of white dwarfs and even of large planets, in particular also for brown dwarfs.

The simplest statistical treatment of atomic structure is the Thomas–Fermi method. Here, one assumes that within each Wigner–Seitz cell, the electrons move in a spherically symmetric potential $V(r)$, which is a solution of the Poisson equation [15].

Inverse Beta-Decay

Another correction enters at high densities due to the inverse beta-decay



Protons and neutrons are generally bound in nuclei. This process plays some role in the neutronization at high densities. Stellar evolution predicts that white dwarfs are formed in stars where the temperature is never high enough to burn much beyond carbon and oxygen, so that massive white dwarfs probably consist of carbon and oxygen.

What happens if the density is increased beyond $8 \times 10^6 \text{ g cm}^{-3}$? At a density of $1.14 \times 10^9 \text{ g cm}^{-3}$, the Fermi energy of the electrons is high enough for the threshold of the inverse beta-decay reaction



The odd–even nucleus immediately undergoes a further electron capture



Cr is then stable to further electron capture until much higher densities. This phase transition softens the EoS, the electrons combine with *Fe* nuclei, and this reduces the pressure.

In C/O white dwarfs the reaction chain $^{12}_6\text{C} \rightarrow ^{12}_5\text{B} \rightarrow ^{12}_4\text{Be}$ has a neutronization threshold of 13.370 MeV, corresponding to a density of $3.9 \times 10^{10} \text{ g cm}^{-3}$. Similar processes occur in Mg/Ne white dwarfs at somewhat lower densities (for more details, see [15]).

5.4 Structure of White Dwarfs and the Chandrasekhar Mass

Since electrons move relativistically inside massive white dwarfs, relativistic effects are expected to play some role in the overall structure of white dwarfs. The correction factors in the TOV equations (Sect. 4.6) are small throughout white dwarfs, so

that we can use the Newtonian approximation for the calculation of white dwarf structure

$$ds_N^2 = - \left(1 - 2 \frac{GM_*}{c^2 r} \right) c^2 dt^2 + dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2). \quad (5.42)$$

In this limit, three-space is flat, i.e. $\exp(\lambda) \equiv 1$ for all radii, and $\Phi(r) \equiv 1 + \Phi(r)/c^2$ with the following structure equations

$$\frac{dM(r)}{dr} = 4\pi\varrho_0(r) r^2 \quad (5.43)$$

$$\frac{dP(r)}{dr} = -\frac{GM(r)\varrho_0(r)}{r^2} \quad (5.44)$$

$$\frac{d\Phi(r)}{dr} = \frac{GM(r)}{r^2}. \quad (5.45)$$

5.4.1 Polytropic Approximation

As we have seen in the discussion of the EoS, there are two limiting cases which can be dealt with in a polytropic approximation, the nonrelativistic limit for the electrons and the extreme relativistic limit at high densities, i.e. at densities much above 10^6 g cm $^{-3}$. Equilibrium configurations with a polytropic EoS of the form $P = K\varrho_0^\Gamma$ can easily be differentiated to yield

$$\frac{1}{r^2} \frac{d}{dr} \left(\frac{r^2}{\varrho} \frac{dP}{dr} \right) = -4\pi G \varrho. \quad (5.46)$$

For the parametrization $\Gamma = 1 + 1/n$ we now introduce dimensionless variables

$$\varrho = \varrho_c \theta^n, \quad r = a \xi \quad (5.47)$$

with $\varrho_c = \varrho(0)$ as the central density and a

$$a = \sqrt{\frac{(n+1)K\varrho_c^{1/n-1}}{4\pi G}}. \quad (5.48)$$

The hydrostatic equilibrium therefore satisfies the following equation (Lané–Emden equation)

$$\frac{1}{\xi^2} \frac{d}{d\xi} \left(\xi^2 \frac{d\theta}{d\xi} \right) = -\theta^n. \quad (5.49)$$

This equation is easily solved on the computer for the initial conditions

$$\theta(0) = 1, \quad \theta'(0) = 0. \quad (5.50)$$

The radius of a star is located at $\theta(\xi_*) = 0$. This requires for $\Gamma = 5/3, n = 3/2$

$$\xi_* = 3.6537 \quad , \quad \xi_*^2 |\theta'(\xi_*)| = 2.71406 \quad (5.51)$$

and for $\Gamma = 4/3, n = 3$

$$\xi_* = 6.89685 \quad , \quad \xi_*^2 |\theta'(\xi_*)| = 2.01824 . \quad (5.52)$$

With ξ_* we can compute the stellar radius R as a function of the central density

$$R = a\xi_* = \sqrt{\frac{(n+1)K}{4\pi G}} \varrho_c^{\frac{1-n}{2n}} \xi_* , \quad (5.53)$$

and with the second quantity the mass M as a function of the central density

$$\begin{aligned} M &= \int_0^R 4\pi r^2 \varrho dr = 4\pi a^3 \varrho_c \int_0^{\xi_*} \xi^2 \theta^n d\xi \\ &= -4\pi a^3 \varrho_c \int_0^{\xi_*} \frac{d}{d\xi} \left(\xi^2 \frac{d\theta}{d\xi} \right) d\xi \\ &= 4\pi a^3 \varrho_c \xi_*^2 |\theta'(\xi_*)| \\ &= 4\pi \left(\frac{(n+1)K}{4\pi G} \right)^{3/2} \varrho_c^{\frac{3-n}{2n}} \xi_*^2 |\theta'(\xi_*)| . \end{aligned} \quad (5.54)$$

We then eliminate the central density and obtain the mass-radius relation for degenerate stars in the polytropic approximation

$$M(R) = 4\pi R^{\frac{3-n}{1-n}} \left(\frac{(n+1)K}{4\pi G} \right)^{\frac{n}{n-1}} \xi_*^2 |\theta'(\xi_*)| \xi_*^{\frac{3-n}{1-n}} . \quad (5.55)$$

With the particular values for the constant K the following exact relations are obtained: for $\Gamma = 5/3$

$$R = 1.12 \times 10^4 \text{ km} \left(\frac{\varrho_c}{10^6 \text{ g cm}^{-3}} \right)^{-1/6} \left(\frac{Y_e}{0.5} \right)^{5/6} \quad (5.56)$$

$$M = 0.496 M_\odot \left(\frac{\varrho_c}{10^6 \text{ g cm}^{-3}} \right)^{1/2} \left(\frac{Y_e}{0.5} \right)^{5/2} \quad (5.57)$$

$$M = 0.70 M_\odot \left(\frac{R}{10^4 \text{ km}} \right)^{-3} \left(\frac{Y_e}{0.5} \right)^5 , \quad (5.58)$$

and for $\Gamma = 4/3$

$$R = 3.347 \times 10^4 \text{ km} \left(\frac{\varrho_c}{10^6 \text{ g cm}^{-3}} \right)^{-1/3} \left(\frac{Y_e}{0.5} \right)^{2/3} \quad (5.59)$$

$$M = M_{\text{Ch}} = 1.457 M_\odot \left(\frac{Y_e}{0.5} \right)^2 . \quad (5.60)$$

For relativistic electrons, the mass is independent of the central density, and therefore also independent of the radius. This is the famous **Chandrasekhar mass** $M_{\text{Ch}} = 1.46 M_\odot$ for $Y_e = 0.5$. It is however smaller for different chemical compositions.

5.4.2 Beyond the Chandrasekhar Treatment

Numerically, the hydrostatic equilibrium can easily be integrated for an algebraic form of the equation of state, either for the ideal EoS equation (5.12) (Fig. 5.15), or for some modified EoS including, e.g. electrostatic corrections (see Hamada and Salpeter [190]).

In such calculations we can use the central density ρ_c as the key parameter in order to obtain the mass as a function of density $M(\rho_c)$ and the radius $R(\rho_c)$. The surface of the object follows for vanishing pressure (or density). In this way we can construct a mass–radius relation $M = M(R)$ (Fig. 5.16) for white dwarf models.

5.4.3 Comparison with Observations

In previous times, the uncertainties in the observed parameters did not allow us to distinguish between C/O and Fe white dwarfs, except for Sirius B. The Hipparcos mission has provided an opportunity of making parallax measurements of hitherto unsurpassed accuracy. Since several of the brightest white dwarfs are included in the Hipparcos catalogue (Table 5.1), this data yields distance measurements from which mass and radius can than then be estimated more reliably than before.

The best direct test of stellar degeneracy is the determination of radii for white dwarfs in visual binaries. In these cases, white dwarf masses follow from their orbital parameters, and stellar radii are derived from a knowledge of the effective temperatures and distances. A knowledge of the mass is lacking for single stars. However, visual binaries with well determined orbital parameters are rare. The observational support for stellar degeneracy rests on four objects in Table 5.2: 40 Eri B, Stein 2051 B, Sirius B and Procyon B.

Prior to Hipparcos, there were only four points in the mass–radius relation test (Fig. 5.17). General relativity introduced gravitational redshift measurements as a second method to determine white dwarf masses without using the mass–radius relation. These gravitational redshift measurements are currently limited by our understanding hydrogen atmospheres (DA). Since this technique uses the exact knowledge of the white dwarf's physical motion to distinguish gravitational velocities from the Doppler effect, a common approach is to use white dwarfs in wide binaries or common proper motion (CPM) pairs in which the system velocity can be accurately determined from the companion. As with surface gravity, gravitational redshift velocity is a function both of mass and radius, requiring either an independent radius determination or an assumed relation between mass and radius. Stellar radii of nearby stars can be derived using effective temperatures and distances (see Table 5.3).

The Hipparcos input catalog contained the visual binaries Sirius, Procyon and 40 Eri B (Provencal et al. [333]). Calculation of the mass employs Kepler's third law. The radii follow from stellar fluxes and distances

$$f_\lambda = 4\pi H_\lambda \frac{R^2}{D^2}. \quad (5.61)$$

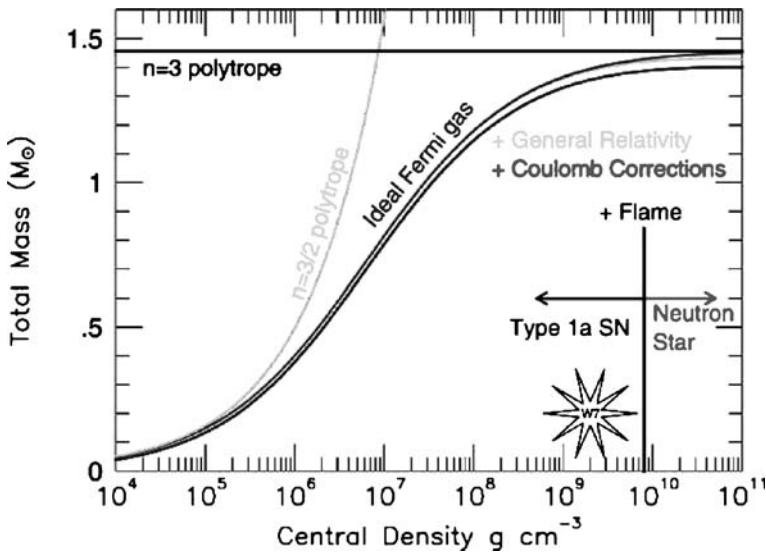


Fig. 5.15. White dwarf masses as a function of central density for the ideal Fermi gas and Salpeter–Hamada model including Coulomb corrections

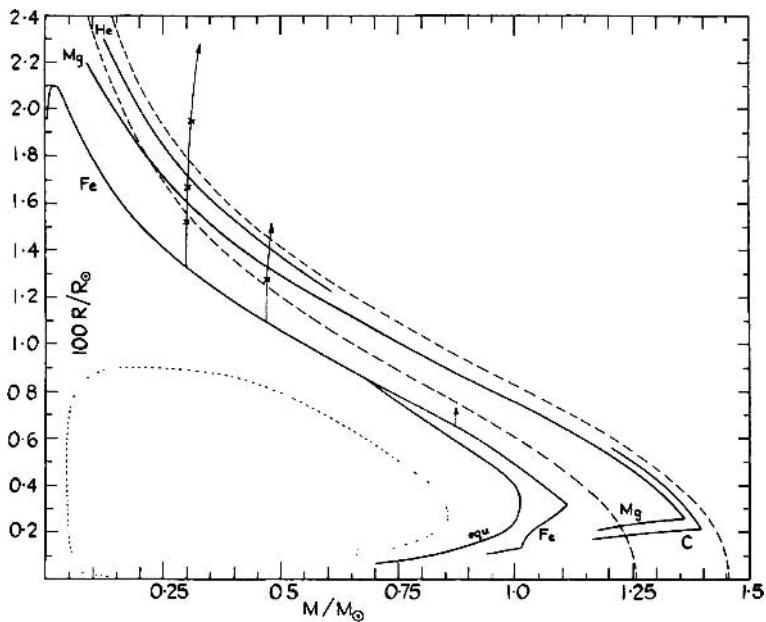


Fig. 5.16. Mass–radius relations for ideal electron gas (*dashed curves*) and the electrostatically corrected form (*solid lines*) (adapted from Hamada and Salpeter [190]). Radii are given in units of $0.01 R_{\odot}$

Table 5.2. White dwarfs in wide binary systems – Hipparcos data (Provencal et al. [333]). The labels G and H refer to ground-based and Hipparcos parallaxes, respectively

Object	π (mas)	Mass (M_\odot)	Radius (R_\odot)	P (yr)
Sirius B	375.6 ± 3.0 (G) 379.2 ± 1.6 (H)	1.03 ± 0.0015	0.0074 ± 0.0007	50.09
Stein 2051B	1812.2 ± 1.1 (G) 1814.4 ± 3.7 (H)	0.48 ± 0.045	0.0111 ± 0.0015	> 300
40 Eri B	208.4 ± 2.3 (G) 198.24 ± 0.84 (H)	0.43 ± 0.02	0.0124 ± 0.0005	> 300
Procyon B	286.4 ± 2.3 (G) 285.9 ± 0.9 (H)	0.602 ± 0.015	0.01234 ± 0.00032	40.5

Table 5.3. Masses, radii and surface temperature of white dwarfs. Data adopted from Provencal et al. [333, 334]

Star	M/M_\odot	R/R_\odot	T [K]
Standard white dwarfs			
Sirius B	1.0034 ± 0.026	0.00840 ± 0.00025	24700 ± 300
G226–29	0.750 ± 0.030	0.01040 ± 0.0003	12000 ± 300
G93–48	0.750 ± 0.060	0.01410 ± 0.0020	18300 ± 300
CD –38 10980	0.740 ± 0.040	0.01245 ± 0.0004	24000 ± 200
L268–92	0.700 ± 0.120	0.01490 ± 0.0010	11800 ± 1000
Stein 2051B	0.660 ± 0.040	0.0110 ± 0.0010	7100 ± 50
Procyon B	0.602 ± 0.015	0.01234 ± 0.00032	7740 ± 50
Wolf 485 A	0.590 ± 0.040	0.01500 ± 0.0010	14100 ± 400
L711–10	0.540 ± 0.040	0.01320 ± 0.0010	19900 ± 400
L481–60	0.530 ± 0.050	0.01200 ± 0.0040	11300 ± 300
40 Eri B	0.501 ± 0.011	0.01360 ± 0.0002	16700 ± 300
G154–B5B	0.460 ± 0.080	0.01300 ± 0.0020	14000 ± 400
Wolf 1346	0.440 ± 0.010	0.01342 ± 0.0006	20000 ± 300
Feige 22	0.410 ± 0.030	0.01367 ± 0.0020	19100 ± 400
Compact dwarfs			
GD 140	0.790 ± 0.020	0.00854 ± 0.0005	21700 ± 300
G156–64	0.590 ± 0.060	0.01100 ± 0.0010	7160 ± 200
EG 21	0.580 ± 0.050	0.01150 ± 0.0004	16200 ± 300
EG 50	0.500 ± 0.020	0.01040 ± 0.0006	21000 ± 300
G181–B5B	0.500 ± 0.050	0.01100 ± 0.0010	13600 ± 500
GD 279	0.440 ± 0.020	0.01290 ± 0.0008	13500 ± 200
WD2007–303	0.440 ± 0.050	0.01280 ± 0.0010	15200 ± 700
G238–44	0.420 ± 0.010	0.01200 ± 0.0010	20200 ± 400

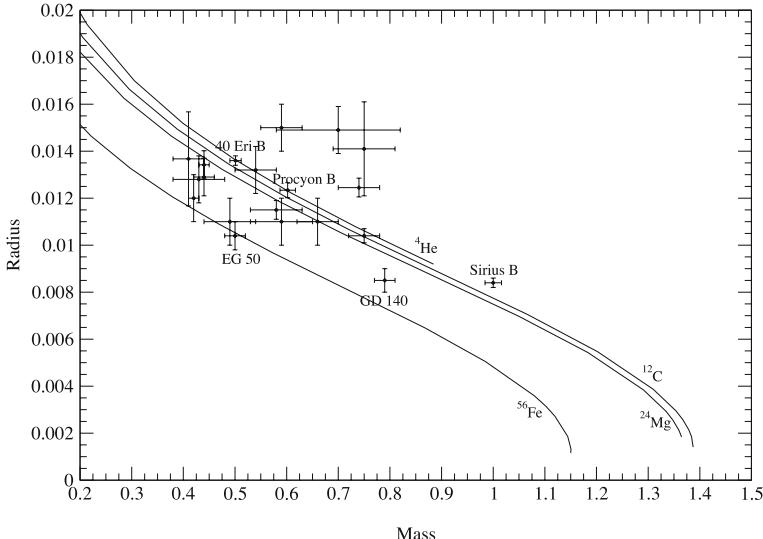


Fig. 5.17. White dwarf masses and radii in solar units derived from Hipparcos data (see Table 5.3). 40 Eri B and Procyon B now moved onto the C/O track. Various white dwarfs still fall onto the iron track – a result which does not follow from standard stellar evolution theory

Gravitational redshift velocity is a function of mass and radius

$$v_g = 63.5 \frac{M}{M_\odot} \frac{0.01 R_\odot}{R} \text{ km s}^{-1}. \quad (5.62)$$

The Hipparcos input catalog includes seven CPM pairs, in which the white dwarf has a gravitational redshift velocity determination.

The mass–radius relation is now more firmly supported by observational grounds (Fig. 5.17). Sirius B and 40 Eri B fit the theoretical line quite precisely. But many stars (such as EG 50 and GD 140) all lie significantly below the theoretical line expected for carbon–oxygen white dwarfs. These objects have radii significantly smaller than predicted. They might contain an iron-rich core; this is, however, not expected on grounds of stellar evolution.

The Hipparcos parallaxes discussed above have enabled to significantly improve the mass and radius determination of some WDs, thus allowing for a direct confrontation with the predictions of WD theory. In particular, the suspicion that some WDs would fall on the zero-temperature, mass–radius relation consistent with iron cores has been placed on a firm observational ground by these satellite-based measurements. Indeed, some WDs have much smaller radii than expected if their interior were made of carbon and oxygen, suggesting that, at least, two of the observed WDs have iron-rich cores. Specifically, the present determinations indicate that GD 140 and EG 50 have radii and masses consistent with zero-temperature, iron WDs. Obviously, such results are in strong contradiction with the standard predictions of stellar

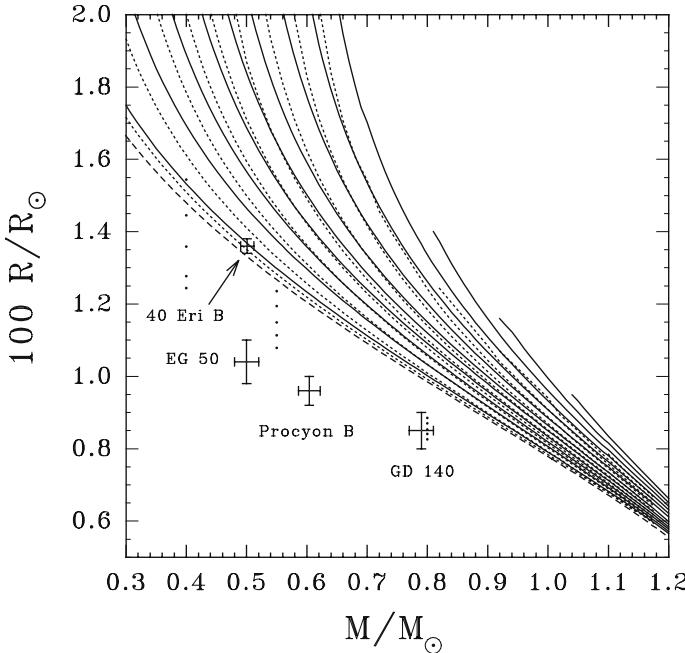


Fig. 5.18. The mass–radius relation for WD stars with a carbon core, surrounded by a helium layer with a thickness of $10^{-2} M_*$. Solid and short dashed lines correspond to the cases of objects with an outermost pure hydrogen layer of $10^{-5} M_*$ and to models without hydrogen layer, respectively. Medium dashed line corresponds to the mass–radius relationship for homogeneous HS carbon models. Included in this figure are values corresponding to T_{eff} (in units of 10^3 K) of 5, 15, 25, 35, 45, 55, 70, 85, 100, 115, 130 and 145. Figure adapted from [324]

evolutionary calculations, which allow for an iron-rich interior only in the case of presupernova objects.

Panei et al. [324] presented accurate and detailed mass–radius relations for white dwarf models with helium, carbon, oxygen, silicon and iron cores, and with and without a hydrogen envelope, by using a fully updated stellar evolutionary code. They considered masses from $0.15 M_\odot$ to $0.5 M_\odot$ for the case of helium core, from 0.45 to $1.2 M_\odot$ for carbon, oxygen, silicon and iron cores. In addition, they explore the effects of gravitational, chemical and thermal diffusion on low mass helium white dwarf models with hydrogen and helium envelopes. In Figs. 5.18 and 5.19, the results for carbon and iron interiors are shown. Although the effects due to finite temperature and the presence of an outer hydrogen envelope are also noticeable, these are not so large as in the case of the low-mass helium WD models. For example, for $1.2 M_\odot$ models, both effects are able to inflate the star only up to $\leq 19\%$. This is expected, because as mass increases, internal density (and electron chemical potential μ_e) also increases. Thus, as thermal effects enter the EoS of the degenerate gas as a correction $\propto (T/\mu_e)^2$, EoS gets closer to the zero temperature

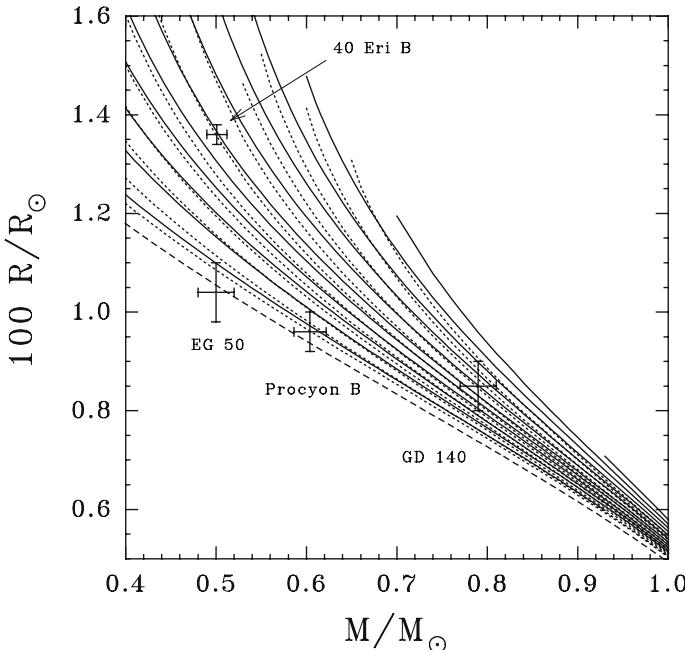


Fig. 5.19. Same as Fig. 5.18, but for an iron core. Included are the Hipparcos data for 40 Eri B, EG 50, Procyon B and GD 140. Figure adapted from [324]

behavior, i.e. to the HS structure. As the thickness of the hydrogen layer is $\propto g^{-1}$ (g is the surface gravity), it also tends to zero for very massive models.

In the case of an iron core, for a fixed mass value, the mean density is almost twice the corresponding to carbon and oxygen cores. Thus, it is not surprising that, for the range of the effective temperature considered here, thermal effects are less important than in the standard case. For example, for the $0.45 M_\odot$ iron model at $T_{\text{eff}} \approx 25,000$ K, thermal effects inflate its radius only by $\approx 17\%$. For the iron core, models are only considered up to a mass value of $1.0 M_\odot$. Higher mass objects are very near the critical mass limit for such a composition (i.e. the central density is at about the neutronization threshold). The evolution towards iron core white dwarfs is, however, still unclear.

5.5 The Relativistic Instability of White Dwarf Stars

Solutions to the TOV equations correspond to stellar configurations that are in hydrostatic equilibrium. Equilibrium does not assure stability, however. Equilibrium configurations may correspond either to a maximum or minimum in the total energy with respect to compression or dilation. In certain ranges of the parameter space there might be two solutions, one at higher density than the other one.

In order to test stability, one has to look at perturbations of the equilibrium solution. Vibrations may either cause their disassembly or their collapse. We now discuss a necessary condition for stability.

5.5.1 Necessary Condition for Stability

One can prove that in a sequence of solutions to the TOV equation for perfect fluids a star can pass from stability to instability with respect to any mode of radial oscillations only at the value of the central density at which the equilibrium mass is stationary, i.e. whenever

$$\frac{\partial M(\varrho_c)}{\partial \varrho_c} = 0. \quad (5.63)$$

In the following we need some relation for the chemical potential μ . From thermodynamics one knows the expression for the pressure P

$$P = -\frac{\partial E}{\partial V} = -\frac{\partial(\varrho/\varrho_0)}{\partial(1/\varrho_0)} = \varrho_0^2 \frac{\partial}{\partial \varrho_0} \left(\frac{\varrho}{\varrho_0} \right) = \varrho_0 \mu - \varrho, \quad (5.64)$$

where $\mu \equiv d\varrho/d\varrho_0$ is the chemical potential. From this relationship between μ and ϱ we find

$$\frac{dP}{d\varrho_0} = \varrho_0 \frac{d\mu}{d\varrho_0}. \quad (5.65)$$

The TOV equation can be written as

$$-\int_r^{r'} d\Phi = \int_r^{r'} \frac{dP}{\varrho c^2 + P} = \int_r^{r'} \frac{d\mu}{\mu} \quad (5.66)$$

for $r < r' < R$. This implies

$$\mu(r) \exp(\Phi(r)) = \mu(r') \exp(\Phi(r')) = \text{const.} \quad (5.67)$$

This relation shows that the baryon chemical potential at any point in a star, corrected by the redshift factor $\alpha(r)$, is a constant. In particular, the chemical potential satisfies

$$\mu(r) \exp(\Phi(r)) = \mu_* \sqrt{1 - \frac{2GM_*}{c^2R_*}}, \quad (5.68)$$

where μ_* is the value at the surface of the star. From this it follows that

$$\frac{\partial M_* c^2}{\partial A} = \mu_* \sqrt{1 - \frac{2GM_*}{c^2R_*}}. \quad (5.69)$$

This can be written as

$$\mu_* \sqrt{1 - \frac{2GM_*}{c^2R_*}} \frac{\partial A}{\partial \varrho_c} = \frac{\partial M_* c^2}{\partial \varrho_c}. \quad (5.70)$$

This proves the statement that an equilibrium star is stationary at the same central densities as the mass.

We can apply this necessary condition for stability to polytropes. *In this case $M(\varrho_c)$ is an increasing function of density for $\Gamma > 4/3$, is stationary for $\Gamma = 4/3$, and is a decreasing function for $\Gamma < 4/3$. We find therefore a necessary condition for stability $\Gamma > 4/3$.*

5.5.2 The Total Energy in the Post-Newtonian Limit

The total energy content of a Newtonian star is given by its internal energy content U and the gravitational potential energy W

$$E_N = U + W = \int_0^R \varrho_0 e dV - \int_0^R \frac{GM(r)}{r} dM(r), \quad (5.71)$$

where $dM = \varrho_0 dV$ is the mass contained in a shell at radius r . For a polytropic equation of state, $P = K\varrho_0^\Gamma$, this can easily be calculated, since the internal energy density then follows from the pressure (via the first thermodynamical law)

$$\varrho_0 e = \frac{P}{\Gamma - 1}. \quad (5.72)$$

The virial theorem tells us then that

$$W = -3(\Gamma - 1) U \quad (5.73)$$

and therefore for the total energy

$$E_N = -\frac{3\Gamma - 4}{3(\Gamma - 1)} |W|. \quad (5.74)$$

For $\Gamma = 5/3$ this gives $U = -W/2$ and $E_N = W/2$, a classical result following from the virial theorem. For a relativistic gas, $\Gamma = 4/3$, the total energy vanishes indicating a critical behavior under this condition. Since the Newtonian energy is just the lowest approximation for the relativistic energy, one has to go to higher order terms in order to check stability or instability.

The total energy of a compact star, excluding rest-mass energy, follows from the total mass M

$$E = (M - M_0) c^2 = c^2 \int_0^R \left[\varrho \sqrt{1 - \frac{2GM(r)}{c^2r}} - \varrho_0 \right] dV \quad (5.75)$$

with the relativistic volume element

$$dV = \frac{1}{\sqrt{1 - GM(r)/c^2r}} 4\pi r^2 dr. \quad (5.76)$$

Replacing the total energy by $\varrho = \varrho_0(1+e/c^2)$ and expanding in the small parameters e and $m(r) \equiv GM(r)/c^2$, we find to second order (in the following $c = 1$)

$$E_{PN} = \int_0^R \varrho_0 \left[e - \frac{m}{r} - e \frac{m}{r} - \frac{1}{2} \left(\frac{m}{r} \right)^2 \right] dV. \quad (5.77)$$

Since $dm' = \varrho_0 dV$ is an invariant, we write a somewhat modified Newtonian part

$$E_N = U + W = \int_0^R \varrho_0 e dV - \int_0^R \frac{m'}{r} dm', \quad (5.78)$$

with $dm' \equiv \varrho_0 dV$. In this way we find an expansion for the total energy

$$E_{PN} = E_N + \Delta E_{GR} \quad (5.79)$$

with

$$\Delta E_{GR} = \int_0^R \varrho_0 dV \left[-e \frac{m}{r} + \frac{1}{2} \left(\frac{m}{r} \right)^2 + \frac{m'}{r} - \frac{m}{r} \right]. \quad (5.80)$$

On the other hand one obtains in first order for the volume

$$V = \frac{4\pi r^3}{3} \left(1 + \frac{3}{r^3} \int_0^r mr dr \right) \quad (5.81)$$

i.e.

$$r' - r = \frac{1}{r^2} \int_0^r mr dr. \quad (5.82)$$

Working to first order in the mass we obtain

$$m'(V') - m(V) = \int_0^V dV \left[\varrho_0 - \varrho \sqrt{1 - \frac{2m}{r}} \right] \simeq - \int_0^V \varrho_0 dV \left(e - \frac{m}{r} \right). \quad (5.83)$$

Therefore, we can write

$$\frac{m'}{r'} - \frac{m}{r} = \frac{m' - m}{r'} - \frac{m(r' - r)}{rr'}. \quad (5.84)$$

Thus, the GR correction can be expressed in terms of five integrals

$$\Delta E_{GR} = I_1 + I_2 + I_3 + I_4 + I_5 \quad (5.85)$$

with the following definitions

$$I_1 = \int_0^M e \frac{m}{r} dm \quad (5.86)$$

$$I_2 = -\frac{1}{2} \int_0^M \left(\frac{m}{r}\right)^2 dm \quad (5.87)$$

$$I_3 = - \int_0^M \frac{dm}{r} \int_0^m e dm \quad (5.88)$$

$$I_4 = \int_0^M \frac{dm}{r} \int_0^m \frac{m}{r} dm \quad (5.89)$$

$$I_5 = - \int_0^M \frac{m dm}{r^4} \int_0^m rm dm . \quad (5.90)$$

In the case of a polytropic equation of state these integrals can be reduced to two of them

$$I_5 = \frac{1}{n} I_1 \quad (5.91)$$

$$I_4 = 2I_2 - \frac{2}{n} I_1 - \frac{3}{n} I_3 \quad (5.92)$$

$$I_3 = I_1 - \frac{2n}{n+1} (I_2 + I_4) . \quad (5.93)$$

Combining these together we find

$$\Delta E_{\text{GR}} = \frac{5+2n-n^2}{n(5-n)} 2I_1 + \frac{n-1}{5-n} 3I_2 . \quad (5.94)$$

Using polytropic expansions we get

$$\Delta E_{\text{GR}} = -0.91829 M^{7/3} \varrho_c^{2/3} . \quad (5.95)$$

5.5.3 GR White Dwarf Instability

The post-Newtonian contributions to the total energy discussed above have a severe consequence for the stability of white dwarfs. This is relevant only for the highest masses possible, i.e. near the Chandrasekhar mass. For this reason, we can concentrate on the treatment of a relativistic Fermi gas, i.e. a polytropic equation of state with $n = 3$, $\Gamma = 1 + 1/n$. As we have seen we can write the total energy in the form

$$E = E_{\text{int}} + E_{\text{grav}} + \Delta E_{\text{int}} + \Delta E_{\text{GR}} . \quad (5.96)$$

In the lowest approximation, only the two first terms are present. We evaluate them for a polytropic equation of state

$$\begin{aligned} E_{\text{int}} &= \int e dm = \int \frac{n P}{\varrho} dm \\ &= K \varrho_c^{1/n} M \frac{n}{|\xi_1^2 \theta'|} \int_0^{\xi_1} \xi^2 \theta'^{n+1} d\xi \end{aligned} \quad (5.97)$$

$$E_{\text{grav}} = -G \int \frac{m}{r} dm = (4\pi \varrho_c)^{1/3} \frac{GM^{5/3}}{|\xi_1^2 \theta'|^{5/3}} \int_0^{\xi_1} \xi^3 \theta' \theta^n d\xi. \quad (5.98)$$

This last integral can be reduced to

$$\int_0^{\xi_1} \xi^3 \theta' \theta^n d\xi = \frac{1}{n+1} \int_0^{\xi_1} \xi^3 \frac{d}{d\xi} \theta'^{n+1} d\xi = -\frac{3}{n+1} \int_0^{\xi_1} \xi^2 \theta'^{n+1} d\xi. \quad (5.99)$$

We have already calculated the last integral in deriving the virial theorem for a ploytropic

$$E_{\text{grav}} = -\frac{3}{5-n} \frac{GM^2}{R}. \quad (5.100)$$

One can show that

$$\frac{M}{R^3} = \frac{4\pi \varrho_c |\theta'|}{\xi_1} \quad (5.101)$$

and therefore

$$E_{\text{grav}} = -\frac{3}{5-n} GM^{5/3} \varrho_c^{1/3} \left| \frac{4\pi \theta'}{\xi_1} \right|^{1/3}. \quad (5.102)$$

Comparing the two expressions for the gravitational energy, we obtain

$$\int_0^{\xi_1} \xi^2 \theta'^{n+1} d\xi = \frac{n+1}{5-n} \xi_1^3 |\theta'|^2. \quad (5.103)$$

Thus we get the lowest order expressions for the total energy

$$E_{\text{int}} = k_1 K \varrho_c^{1/n} M \quad (5.104)$$

$$E_{\text{grav}} = -k_2 G \varrho_c^{1/3} M^{5/3} \quad (5.105)$$

with the two constants for $n = 3$

$$k_1 = \frac{n(n+1)}{5-n} \frac{|\xi_1^2 \theta'|}{\xi_1} = 1.75579 \quad (5.106)$$

$$k_2 = \frac{3}{5-n} \frac{|4\pi \xi_1^2 \theta'|^{1/3}}{\xi_1} = 0.639001. \quad (5.107)$$

The term ΔE_{int} represents the changes of the equation of state from that of an exact polytrope $n = 3$, since the electrons are not completely relativistic. The internal energy per unit mass is

$$e = \frac{\epsilon_e - n_e m_e c^2}{\varrho_0}, \quad (5.108)$$

where

$$\varrho_0 = \frac{m_u n_e}{Y_e}. \quad (5.109)$$

Using the expansion for the extreme relativistic limit

$$e = \frac{3}{4} \frac{Y_e m_e c^2}{m_u} \left(x - \frac{4}{3} + \frac{1}{x} + \dots \right), \quad (5.110)$$

the term proportional to x is simply $3P/\varrho_0$ and has already been used to calculate E_{int} . The next term is a constant and can be dropped for variational calculations. Therefore, the dominant term remains

$$\Delta E_{\text{int}} = \frac{3}{4} \frac{Y_e m_e c^2}{m_u} \int \frac{1}{x} dm. \quad (5.111)$$

Since x is given as

$$x = \left(\frac{3\pi^2 Y_e \varrho_0 A_e^3}{m_u} \right)^{1/3}, \quad (5.112)$$

we can evaluate the integral for a polytrope $n = 3$ with errors being of higher order

$$\Delta E_{\text{int}} = k_3 \frac{m_e c^3}{\hbar (m_u / Y_e)^{2/3}} M \varrho_c^{-1/3}. \quad (5.113)$$

The constant is given as

$$k_3 = \frac{3}{4} \frac{1}{(3\pi^2)^{1/3}} \frac{1}{|\xi_1^2 \theta'|} \int_0^{\xi_1} \xi^2 \theta^2 d\xi = 0.51972. \quad (5.114)$$

The general relativistic correction has been calculated in (5.95)

$$\Delta E_{GR} = -k_4 \frac{G^2}{c^2} M^{7/3} \varrho_c^{2/3} \quad (5.115)$$

with

$$k_4 = 0.91829. \quad (5.116)$$

Thus the total energy in post-Newtonian order can be summarized as follows

$$E_{PN} = (AM - BM^{5/3}) \varrho_c^{1/3} + C M \varrho_c^{-1/3} - D M^{7/3} \varrho_c^{2/3} \quad (5.117)$$

with the following constants

$$A = k_1 K \quad , \quad B = k_2 G \quad , \quad C = k_3 \frac{m_e c^3}{\hbar(m_u/Y_e)^{2/3}} \quad , \quad D = k_4 \frac{G^2}{c^2}. \quad (5.118)$$

Equilibrium is achieved for $\partial E/\partial \varrho_c = 0$, i.e. for

$$0 = \frac{1}{3} (AM - BM^{5/3}) \varrho_c^{-2/3} - \frac{1}{3} C M \varrho_c^{-4/3} - \frac{2}{3} D M^{7/3} \varrho_c^{-1/3}. \quad (5.119)$$

To leading orders we can neglect the post-Newtonian corrections and obtain the Chandrasekhar mass

$$M_\infty = \left(\frac{A}{B} \right)^{3/2} = 1.457 \left(\frac{Y_e}{0.5} \right)^2 M_\odot. \quad (5.120)$$

The **onset of instability** occurs when $\partial^2 E/\partial \varrho_c^2 = 0$, i.e. when

$$0 = -\frac{2}{9} (AM - BM^{5/3}) \varrho_c^{-5/3} + \frac{4}{9} C M \varrho_c^{-7/3} + \frac{2}{9} D M^{7/3} \varrho_c^{-4/3}. \quad (5.121)$$

We now solve for $AM - BM^{5/3}$ from the equilibrium condition and substitute this into the above condition and replace therein $M = (A/B)^{3/2}$, since all terms are of the same order of magnitude, to obtain a critical density for instability

$$\varrho_c^{\text{crit}} = \frac{CB^2}{DA^2} = \frac{16k_3 k_2^2}{(3\pi^2)^{2/3} k_4 k_1^2} \frac{m_u^2}{Y_e^2 m_e A_e^3}, \quad (5.122)$$

or numerically

$$\varrho_c^{\text{crit}} = 2.65 \times 10^{10} \left(\frac{Y_e}{0.5} \right)^{-2} \text{g cm}^{-3}. \quad (5.123)$$

This is the critical density for **the onset of instability due to general relativity** in a white dwarf. In particular, for Fe we have $\varrho_c^{\text{crit}} = 3.07 \times 10^{10} \text{ g cm}^{-3}$. This is however higher than the inverse beta threshold of $1.4 \times 10^9 \text{ g cm}^{-3}$. GR effects are therefore not important for iron white dwarfs, they could however be important for C/O white dwarfs, where the critical density is somewhat smaller than the neutronization thresholds. In these cases, **relativistic effects limit the central density**.

5.6 Cooling White Dwarfs

When white dwarfs are formed after the AGB phase, their internal temperature is fairly high, of the order of 100 Million K. The interior of the white dwarf is completely degenerate, so that the conductivity is high (i.e. a large mean free path for the electrons). The core of the white dwarf is therefore practically isothermal with a temperature T_c . Cooling of the white dwarf occurs only over the surface of the star which has a temperature T_* . The gaseous layers above the degenerate core act as a thermal shield, because heat must be transported by means of photon diffusion through these surface layers.

5.6.1 Structure of the Surface Layers

The luminosity resulting from photon diffusion across the surface layers is given by

$$L_* = -4\pi r^2 \frac{c}{3\kappa_R \varrho_0} \frac{d}{dr} (aT^4) . \quad (5.124)$$

κ_R denotes the Rosseland mean for the opacity of the gaseous material. For high temperatures, we may use Kramer's opacity, $\kappa = \kappa_0 \varrho_0 T^{7/2}$, so that we get for the temperature gradient

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\kappa \varrho_0}{T^3} \frac{L_*}{4\pi r^2} . \quad (5.125)$$

Dividing this equation by the hydrostatic equilibrium

$$\frac{dP}{dr} = -\frac{GM(r)\varrho_0}{r^2} \quad (5.126)$$

we obtain

$$\frac{dP}{dT} = \frac{4ac}{3} \frac{4\pi GM(r)}{\kappa_0 L_*} \frac{T^{13/2}}{\varrho_0} . \quad (5.127)$$

In the surface layers we can assume that the mass is essentially constant, and the density ϱ_0 can be eliminated in terms of the equation of state for nondegenerate matter

$$P dP = \frac{4ac}{3} \frac{4\pi GM_*}{\kappa_0 L_*} \frac{k_B}{\mu m_u} T^{15/2} dT . \quad (5.128)$$

This can easily be integrated with the boundary conditions $P = 0$ and $T = 0$

$$\varrho_0(r) = \left(\frac{2}{8.5} \frac{4ac}{3} \frac{4\pi GM_*}{\kappa_0 L_*} \frac{\mu m_u}{k_B} \right)^{1/2} T^{3.25}(r) . \quad (5.129)$$

With the appropriate value for κ_0

$$\kappa_0 = 4.34 \times 10^{24} Z(1 + X) \frac{\text{cm}^2}{\text{g}} , \quad (5.130)$$

we get a solution for the density as a function of the temperature in the surface layers.

This approximation will break down when the thermal pressure equals the degenerate pressure, i.e. for

$$\frac{\varrho_0 k_B T}{\mu m_u} = 1.0 \times 10^{13} \left(\frac{\varrho_0}{\mu_e} \right)^{5/3} . \quad (5.131)$$

This gives the density for temperatures in units of 10^6 K as

$$\rho_0 = 24 \text{ g cm}^{-3} \mu_e T_6^{3/2}. \quad (5.132)$$

With our density–temperature relation we obtain then a luminosity

$$L_* = 5.7 \times 10^{26} \text{ erg s}^{-1} \frac{\mu}{\mu_e^2} \frac{1}{Z(1+X)} \frac{M}{M_\odot} T_6^{7/2}. \quad (5.133)$$

Thus given the observed luminosity, we can derive the interior temperature of the white dwarf core. For a typical composition for a He white dwarf, $X = 0$, $Y = 0.9$ and $Z = 0.1$ with $M = M_\odot$, $\mu_e = 2$ and $\mu = 1.4$ we find

$$L_* = 2 \times 10^{27} \text{ erg s}^{-1} \frac{M}{M_\odot} T_6^{7/2}. \quad (5.134)$$

Since observed luminosities are in the range of $(10^{-2}\text{--}10^{-5}) L_\odot$, the core temperatures are at least a few million degrees. A more accurate calculation for the relation between core temperature and luminosity is shown in Fig. 5.20. Advances have been obtained in the last years in the question of conductive and radiative opacities for the envelope, as well as for the detailed description of the thermodynamic properties of the dense fully ionized interior plasma, which contains the effects of crystallization (see, e.g. Chabrier et al. 2000). The break in the temperature–luminosity relation

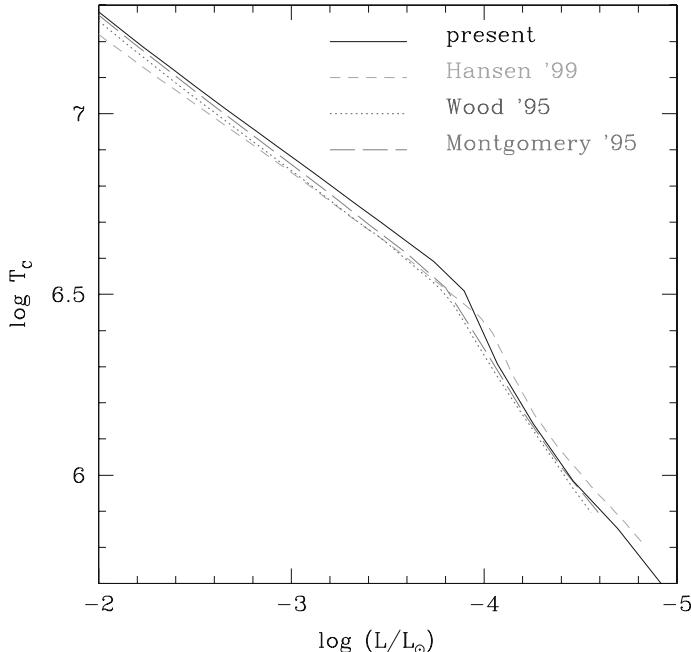


Fig. 5.20. Core temperature vs. luminosity for white dwarfs (adapted from Chabrier et al. [111])

occurs due to the onset of convection for low temperatures in the atmosphere. For luminosities larger than those of the changes in the slope, it has been demonstrated that the core temperature is completely insensitive to the stratification of the upper envelope and the details of the atmosphere.

5.6.2 Cooling Curves and Crystallization

In a fully crystallized white dwarf the only contribution to the luminosity comes from the thermal reservoir of the core and the contraction of the surface layers

$$L_*(T_c) = - \int_0^M C_V \frac{dT}{dt} dm - \text{compressional work}. \quad (5.135)$$

Most of the thermal heat stems from the specific heat C_V of the quantum solid. For a nondegenerate monatomic gas we would have $C_V = 3k_B/2$, and therefore the total thermal energy contained in the core of a white dwarf would be

$$U = \frac{3}{2} k_B T \frac{M_*}{N_B m_u}. \quad (5.136)$$

This energy store is considerable, $U \simeq 10^{48}$ ergs for $T_c = 10^7$ K. At low temperatures, the ions form however a crystal lattice with C_V exceeding somewhat the value for free ions. At extremely low temperatures, zero-point quantum fluctuations will determine the behavior of C_V as a function of temperature

$$C_V = 3k_B D_3(x_D) \quad , \quad D_3(x_D) = \frac{3}{x_D^3} \int_0^{x_D} \frac{x^4 \exp x dx}{(\exp x - 1)^2}, \quad (5.137)$$

where $x_D = T_D/T$ and T_D is the Debye temperature of the lattice (Fig. 5.21). For $T < T_D$, i.e. $x_D > 1$, we can write

$$D_3(x_D) = \frac{12}{x_D^3} \int_0^{x_D} \frac{x^3 dx}{\exp x - 1} = \frac{4\pi^4}{5x_D^3}. \quad (5.138)$$

For low temperatures we find the typical phonon behavior for a quantum lattice

$$C_V = \frac{12\pi^4}{5} k_B \left(\frac{T}{T_D} \right)^3, \quad (5.139)$$

The Debye temperature is usually defined in terms of the plasma frequency Ω_p of the ions

$$\Omega_p = \sqrt{\frac{4\pi n_i Z^2 e^2}{m_i}} \quad (5.140)$$

as the temperature, where the vibrational energy equals the thermal energy,

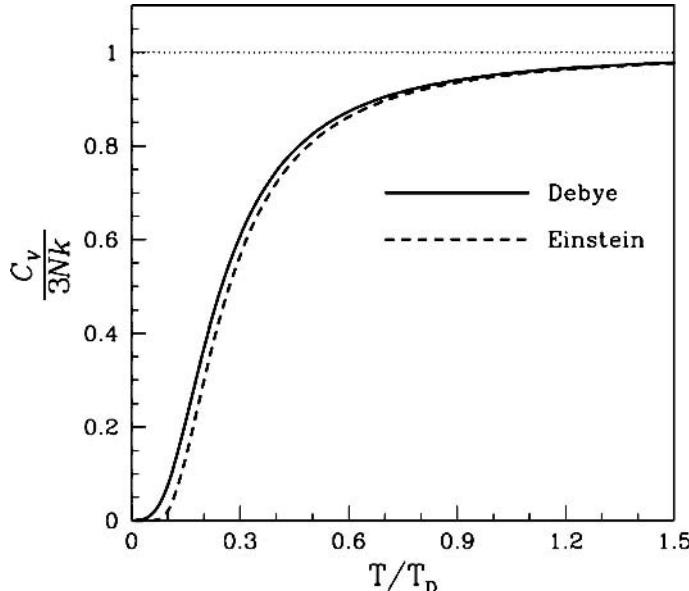


Fig. 5.21. Specific heat capacity of diamond as a function of temperature compared to the Debye and Einstein model. The Debye temperature T_D for diamond is 2200 K under lab conditions

$$k_B T_D = \hbar \Omega_p , \quad (5.141)$$

or numerically for the interior of white dwarfs

$$T_D = 4 \times 10^6 \varrho_{0,6}^{1/2} \text{ K.} \quad (5.142)$$

For white dwarfs with masses in the range of $0.6 M_\odot$ the Debye temperature is in the range of a few million degrees. Another characteristic temperature is T_g , the point at which the ion kinetic energy exceeds its vibrational energy. Above this temperature, the lattice completely dissolves, yielding a dense imperfect gas. The corresponding temperature is

$$T_g \simeq 3 \times 10^6 \varrho_{0,6}^{1/3} Z^{5/3} \text{ K.} \quad (5.143)$$

For a C/O white dwarf, this is at least a factor 10 higher than the Debye temperature. Therefore, in the intermediate range, $T_D < T_c < T_g$, the lattice approximation has to be used (for more details, see [15]).

The integration of (5.135) together with the relation of core temperature vs. luminosity gives then the luminosity as a function of time (Fig. 5.22).

In general one finds that before crystallization sets in, massive white dwarfs evolve slowly in temperature because of their greater energy content and their smaller radiative areas (see mass-radius relation). Since they are hotter and brighter at a given

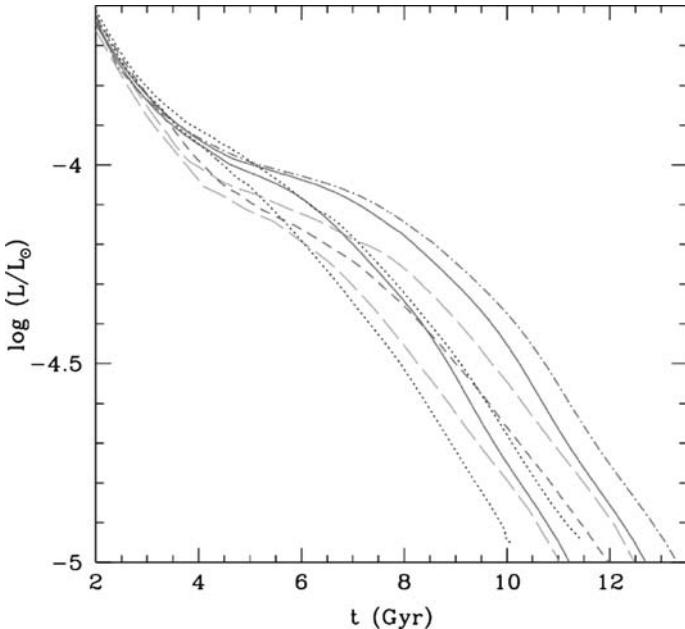


Fig. 5.22. Cooling sequences for a $0.6 M_{\odot}$ WD with hydrogen atmosphere under various model assumptions (Chabrier et al. [111]). Crystallization occurs at about 6 Gyrs

core temperature than less massive white dwarfs, and since crystallization occurs always at the same internal temperature, massive white dwarfs crystallize earlier and at higher T_{eff} and at higher luminosity. At this stage the crystallized core enters the Debye-cooling regime ($C_V \propto T_c^3$) and cools more quickly (see Fig. 5.23).

In the Hubble Deep Field (HDF), one has found faint blue objects which have been interpreted as halo white dwarfs (Ibata et al. [210]). This would be consistent with the interpretation of microlensing events towards the large Magellanic Clouds as stellar remnants. In addition, very cool ($T_{\text{eff}} \leq 4000$ K) white dwarfs with high proper motions have been found in the last years. In Fig. 5.23 the expected absolute magnitudes are shown as a function of time for various masses and various assumptions. White dwarfs with He atmospheres cool much faster and are not the appropriate interpretation for extremely old candidates.

5.6.3 Testing WD Crystallization Theory

The crystallization process leads to one of the biggest sources of uncertainty in the ages of cool white dwarf stars. When a typical mass white dwarf star with mass $\simeq 0.6 M_{\odot}$ cools down to $T_{\text{eff}} \simeq 6000\text{--}8000$ K (depending on the core composition), the high-density core will undergo a phase transition from liquid to solid. As a consequence of this, a latent heat of crystallization will be released, providing a new source of thermal energy that introduces a delay in the gradual cooling of the star

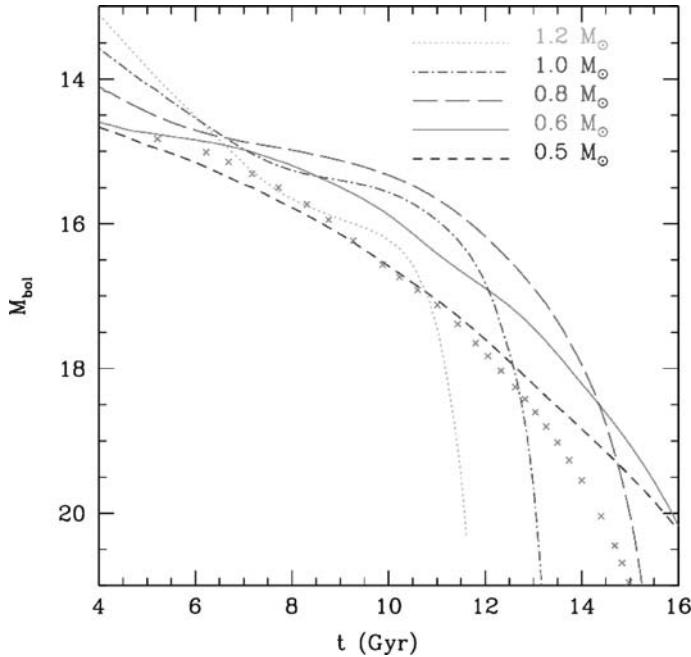


Fig. 5.23. Bolometric magnitudes as a function of time for various cooling WDs parametrized by the mass of the WD. Figure adapted from Chabrier et al. [111]

(Hansen and Liebert [191]). This source of uncertainty can be calibrated by using observations of pulsating white dwarfs (Metcalfe et al. [282]).

5.7 White Dwarfs in Binary Systems

Cataclysmic variables (CVs) are stellar **binary systems** that consist of a normal star and a white dwarf. They are typically small – a typical binary system is roughly the size of the Earth–Moon system – with an orbital period in the range 1–10 hours. The companion star, a more or less normal star like our Sun, loses material onto the white dwarf by accretion.

There are two principal **energy sources** in a cataclysmic variable: accretion and nuclear fusion. Since the white dwarf is very dense, the gravitational potential energy is enormous, and it is converted into X-rays during the accretion process. The efficiency of this process is typically around 0.03% (i.e. the energy released in X-rays is about 0.03% of the energy which would be released by the total annihilation of the same quantity of matter). The fusion of four hydrogen nuclei into a He nucleus has an efficiency of 0.7%. These are much lower than accretion onto neutron stars (10%) or black holes, observed in X-ray binary systems, but still high enough to make CVs much brighter in X-rays than typical stellar coronae.

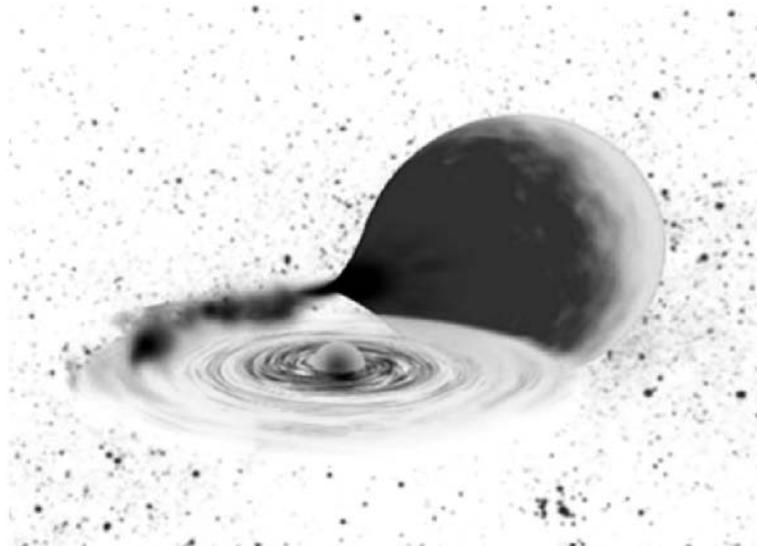


Fig. 5.24. Structure of a cataclysmic system

In the fusion-dominated phase, there are currently two different categories of CVs:

- Classical novae
- Supersoft sources

In the accretion-dominated phase, CVs can be categorized into several different types:

- Dwarf novae
- Polars (AM Her systems)
- Intermediate Polars (DQ Her systems)

Classical Novae

A classical nova occurs when material accreting onto the surface of a white dwarf star's surface begins an unstable thermonuclear fusion reaction. Mass transfer will occur gradually between novae explosions and it will be accreted onto the white dwarf's surface at a low rate, generating only weak X-ray emission. The nova outburst will increase the apparent brightness of the binary star system by 10 thousand to a million times (an increase in stellar magnitude of 10 to 15). They are hypothesized to recur, though with recurrence time-scales of 10,000 years, no known novae recurrence have been observed.

In the X-ray regime, classical novae have been observed many times. There is no clear pattern in their light-curves over a period of years or decades. Several sources have been observed in the period of days to months after the nova explosion to be

weak X-ray emitters (10^{33} – 10^{34} ergs/s), as compared to an Eddington luminosity limit of around 10^{38} ergs/s for a one solar mass white dwarf. One explanation for this weak emission is shock heating of the circumstellar envelope in the white dwarf's environment from new material expelled by the nova explosion. At least two sources (GQ Mus and V1974 Cyg) have been observed to develop a supersoft source-like emission on the time-scales of years. On the other hand, of 26 novae which exploded in the decade prior to the ROSAT All-Sky Survey, only GQ Mus was observed to have become a supersoft source. All other sources surveyed in the All-Sky Survey exhibited only upper limits. Novae that are close enough for sensitive observations are detected in hard X-rays decades after their eruptions; this is presumably due to on-going accretion onto the white dwarf. Examples of these nearby systems with hard X-ray emission include Pup 91, QU Vul, and CP Pup.

Supersoft Sources

Supersoft sources (or SSSs) are a recent addition to the family of cataclysmic variables, first categorized by ROSAT observations although several sources were observed and recognized by prior missions. SSSs are objects with temperatures of between 200,000 and 800,000 K and luminosities around 10^{38} ergs/s. More than 90% of their observed X-ray emission is below 0.5 keV. Currently, it is a subject of debate exactly what physical object is associated with these high X-ray luminosities and low temperatures. There are two leading theories. In the first, SSSs consist of a neutron star or black hole surrounded by a bright and highly extended accretion disk or cloud, which reprocesses energetic photons produced by the accretion onto the compact object. In the other theory, SSSs are white dwarfs with classic hydrogen fusion occurring from material accreting onto their surfaces. The second theory would make SSSs the progenitor for Type Ia supernovae.

In the low mass X-ray binary (LMXB) model of SSSs, the lower temperature spectra of these systems with comparison to their standard LMXB cousins is due to an extended cloud around the neutron star or black hole. This cloud Compton scatters the original higher temperature X-ray photons and the larger observed black-body radius (of the order 10^4 km) is simply the radial size of this cocoon of matter. The compact object would need to have a weak and small magnetic field structure: a strong or extended field would disrupt the cloud. Theoretical calculations of the spectra for spherical accretion onto a nonmagnetized neutron star seem to match observations to date reasonably well.

In the white dwarf scenario, the observed black-body radii is simply the size of the white dwarf star itself with nuclear fusion occurring on its surface. If accretion occurs onto the white dwarf surface at low rates, fusion will be sporadic and violent, resulting in classical novae type explosions. If accretion is at a high rate, the white dwarf will acquire a red-giant-like atmosphere. Continuous nuclear fusion on the dwarf star surface would be possible only for a narrow range of accretion rates of the order of 10^{-7} solar masses per year. For relatively massive white dwarfs (0.7–1.2 solar masses), the luminosities of nuclear fusion and temperature would match the ROSAT observations. This raises the distinct possibility that such objects could

eventually exceed the Chandrasekhar limit, making them the progenitors for Type Ia supernovae.

Dwarf Novae

Dwarf novae outbursts are intrinsically much less luminous events than classical novae outbursts. Their peak absolute magnitudes are at least 100 times weaker. Dwarf novae are known to recur, with some recurring on times as short as a few weeks. Dwarf novae also have short durations, lasting a few days. Dwarf novae also can exhibit a variety of unusual behaviors. SU UMa type sources occasionally exhibit extremely long outbursts known as super-outbursts. Z Cam stars will occasionally get stuck in a standstill during which their brightness is both below outburst stage and well above quiescent levels. VY Scl stars, also known as antidwarf novae, will spend most of their time in an outburst state, with occasional dips into quiescence that last for a few days. Finally, there are nova-like objects which behave much like novae long after their eruptions, but which have never exhibited novae outbursts. They are also distinct from dwarf novae outbursts in that they have permanently high rates of mass transfer.

The principal source of electromagnetic radiation in a dwarf nova system is the accretion disk. The companion star to the white dwarf is a low mass red dwarf star filling its Roche lobe with matter streaming onto the accretion disk through the inner Lagrange point. The gas stream from the L1 point impacts the accretion disk and creates a hot spot (Fig. 5.24). Matter gradually transports through the accretion disk onto the surface of the white dwarf, generating temperatures which make the disk much hotter and brighter than either star. The dwarf nova outburst and other related phenomenon are believed to be caused by variations in the accretion rates through the disk. Material reaching the white dwarf surface through the disk must pass through a violent transition region, called the boundary layer: it is here that the X-rays in dwarf novae originate. This is shown dramatically by the recent observations of X-ray eclipses in HT Cas; the eclipse duration is the same as that of the white dwarf as determined by optical observations. The sharpness of the transitions into and out of the eclipse proves that the X-ray emitting region has a size comparable to that of the white dwarf.

Polars (AM Her Systems)

In a polar system such as the prototype AM Her, matter will overflow the Roche lobe of the companion star. However, the white dwarf possesses a strong magnetic field, which prevents the formation of an accretion disk. Instead, the overflowing material is directed by the magnetic field structure until it impacts on the surface of the white dwarf at its magnetic pole. Until impact, the material essentially free falls, thus reaching substantial velocities which are seen in the optical spectra. The collision generates a shock wave which is the source of hard (energetic) X-rays. Hard X-rays emitted in the direction of the white dwarf from the shock wave above

its surface heat the local area around the pole sufficiently for the pole to become a source of intense soft (less energetic) X-rays. Since soft X-rays are coming only from the pole, rotation of the white dwarf can occult the X-ray source on its surface. Polars are generally much stronger sources of soft X-rays than hard X-rays. Most likely, this is due to uneven matter streaming. Clumps in the accretion flows would most likely cause energy to also be liberated deep within the atmosphere of the white dwarf, resulting in more soft X-ray emission. The strong magnetic field will also lock the orientation of the white dwarf relative to the companion, so that orbital and rotational periods are identical.

X-ray emission from polar systems is entirely due to the accretion column and its impact, so in quiescent times when matter is not accreting onto the system, the entire system is much dimmer. Spectral lines measured at these times show the Zeeman effect which measures the magnetic field strength in the megagauss range.

Intermediate Polars (DQ Her systems)

AM Her type systems are distinct from all other CVs in that they completely lack an accretion disk. However, in systems which have less strong magnetic fields, or wider separations between the companion and white dwarf, an accretion disk can form. As material migrates inwards in the disk, it may eventually encounter a magnetic field strong enough to control the flow of material, at which point matter would instead stream from the inner edge of the disk along magnetic field lines onto the pole of the white dwarf (Fig. 5.25). Such a system would therefore be expected to be a source of hard X-rays from the shock at the magnetic poles. CVs which possess both an accretion disk and magnetic fields which disrupt the inner edge of the disk are known as intermediate polars. Intermediate polars, either due to weaker magnetic fields or wider star separations, will not necessarily have orbital and spin rates locked. Observed systems have longer orbital periods than polars, which given that the systems have comparable masses verifies their wider separation. Whether such systems also have weaker magnetic fields is an area of active research.



Fig. 5.25. Structure of a magnetic intermediate polar cataclysmic system. Material migrates inwards in the disk, it may eventually encounter a magnetic field strong enough to control the flow of material onto the white dwarf surface

Problems

5.1. Newtonian Limit: Explain why

- GR effects are important for neutron stars, but not for white dwarfs;
- inverse beta-decay becomes energetically favorable for densities higher than in white dwarfs.

5.2. Lané–Emden Equation: Integrate numerically the Lané–Emden equation for various indices n and determine the boundary values ξ_* and $\xi_*^2 |\theta'(\xi_*)|$ as a function of n .

5.3. Computer Code: Develop a computer code to solve the stellar structure equations for white dwarfs, in particular for the analytic EoS of degenerate matter. Extend this code to involve the integration of a hydrogen envelope [324].

5.4. White Dwarfs from SDSS: Study the spectroscopic white dwarf and hot subdwarf sample from the SDSS first data release, DR1 [227]. They find 2551 white dwarf stars of various types, 240 hot subdwarf stars, and an additional 144 objects identified as uncertain white dwarf stars. Of the white dwarf stars, 1888 are non-magnetic DA types and 171, nonmagnetic DBs. The remaining (492) objects consist of all different types of white dwarf stars: DO, DQ, DC, DH, DZ, hybrid stars like DAB, etc., and those with nondegenerate companions.

5.5. Accretion onto White Dwarfs: Use standard accretion disk theory (see Sect. 10.3) to estimate the surface temperature of accretion disks around white dwarfs, for typical accretion rates of $\dot{M} \simeq 10^{-9} M_\odot \text{ yr}^{-1}$. Determine the luminosity as a function of the accretion rate.

5.6. Magnetic Accretion: What is the minimal magnetization of a white dwarf to prevent the accretion disk to reach to the surface of a white dwarf? Use for this estimate a dipolar magnetosphere with surface field strength B_* .

6 Neutron Stars

Neutron stars are about 20 km in diameter and have a mass of about 1.4 times that of our Sun. This means that a neutron star is so dense that on Earth, one teaspoonful would weigh a billion tons. Because of its small size and high density, a neutron star possesses a surface gravitational field about 2×10^{11} times that of Earth. Neutron stars can also carry magnetic fields a million times stronger than the strongest magnetic fields produced on Earth.

Neutron stars are one of the possible end states for a massive star. They result from massive stars which have mass greater than 6–8 times that of our Sun. After these stars have finished burning their nuclear fuel, they undergo a supernova explosion. This explosion blows off the outer layers of a star into a beautiful supernova remnant. The central region of the star collapses under gravity. It collapses so much that protons and electrons combine to form neutrons.

Massive stars at the end of their lives are believed to consist of a white dwarf-like iron core of mass $(1.2\text{--}1.4) M_{\odot}$, having low entropy ($s \leq k_B$), and surrounded by layers of less processed material from nuclear shell burning. The effective Chandrasekhar mass is dictated by the lepton number Y_L believed to be around 0.41–0.43. As mass is added to the core by shell Si-burning, the core becomes unstable and collapses.

During the collapse, the lepton content decreases due to net electron capture on nuclei and free protons. But when the core density approaches $10^{12} \text{ g cm}^{-3}$, the neutrinos can no longer escape from the core on the dynamical time-scale. After neutrinos become trapped, the lepton number is frozen at a value of about 0.38–0.40, and the entropy also remains fixed. The core continues to collapse until the rapidly increasing pressure reverses the collapse at a bounce density of a few times nuclear density.

This bounce results in a shock which is largely dissipated by the energy required to dissociate massive nuclei in the still infalling matter of the original iron core. The larger the lepton number Y_L of the core, the larger its mass and the smaller this shell. The final lepton number is then controlled by weak interactions, and is strongly dependent upon the number of protons, x_p . So the properties of nuclear matter determine largely the outcome of the collapse, in particular the resulting mass of the newly formed neutron star. Many questions are still open in this field. Today, several accurate mass determinations for neutron stars are available, and they all lie in the narrow range of $(1.25\text{--}1.44) M_{\odot}$.

Neutron stars may appear in supernova remnants, as isolated objects, or in binary systems. One neutron star has even been found to have planets. When a neutron star

is in a binary system, astronomers are able to measure its mass. From a number of such binaries seen with radio or X-ray telescopes, the neutron star mass has been found to be close to have masses of about 1.4 times the mass of the Sun. For binary systems containing an unknown object, this information helps distinguish whether the object is a neutron star or a black hole, since black holes are more massive than neutron stars.

Pulsars are rotating neutron stars. And pulsars pulse because they rotate. Pulsars were first discovered in late 1967 by graduate student Jocelyn Bell Burnett as radio sources that blink on and off at a constant frequency. Now we observe the brightest ones at almost every wavelength of light. Pulsars are spinning neutron stars that have jets of particles moving almost at the speed of light streaming out above their magnetic poles. These jets produce very powerful beams of light. For a similar reason that “true north” and “magnetic north” are different on Earth, the magnetic and rotational axes of a pulsar are also misaligned. Therefore, the beams of light from the jets sweep around as the pulsar rotates, just as the spotlight in a lighthouse does.

A very different type of pulsar is seen by X-ray telescopes in some X-ray binaries. In this case, a neutron star and a normal star form the binary system. The strong gravitational force from the neutron star pulls material from the normal star. The material is funnelled onto the neutron star at its magnetic poles. In this process, called accretion, the material becomes so hot that it produces X-rays. The pulses of X-rays are seen when the hot spots on the spinning neutron star rotate through our line-of-sight from Earth. These pulsars are sometimes called *accretion-powered pulsars* to distinguish them from the *spin-powered radio pulsars*.

6.1 The Structure of a Neutron Star

The cross-section of a neutron star can roughly be divided into four distinct regions (see Fig. 6.1):

- The *atmosphere* which is only a few cm thick.
- The *outer crust* which consists of a lattice of atomic nuclei and Fermi liquid of relativistic degenerate electrons. This is essentially white dwarf matter.
- The outer crust envelops the *inner crust*, which extends from the neutron drip density to a transition density $\rho_{\text{tr}} \simeq 1.7 \times 10^{14} \text{ g cm}^{-3}$.
- Beyond the transition density one enters the *core*, where all atomic nuclei have been dissolved into their constituents, neutrons and protons. Due to the high Fermi pressure, the core might also contain hyperons, more massive baryon resonances, and possibly a gas of free up, down and strange quarks. Finally, π - and K-meson condensates may be found there too.

The equation of state for the outer and inner crust is well-known and described by the model of BPS [64] and Negele and Vautherin [310].

Today, neutron stars come in various flavors depending on the composition of the core. In this respect, we speak now of traditional neutron stars (or **hadronic stars**),

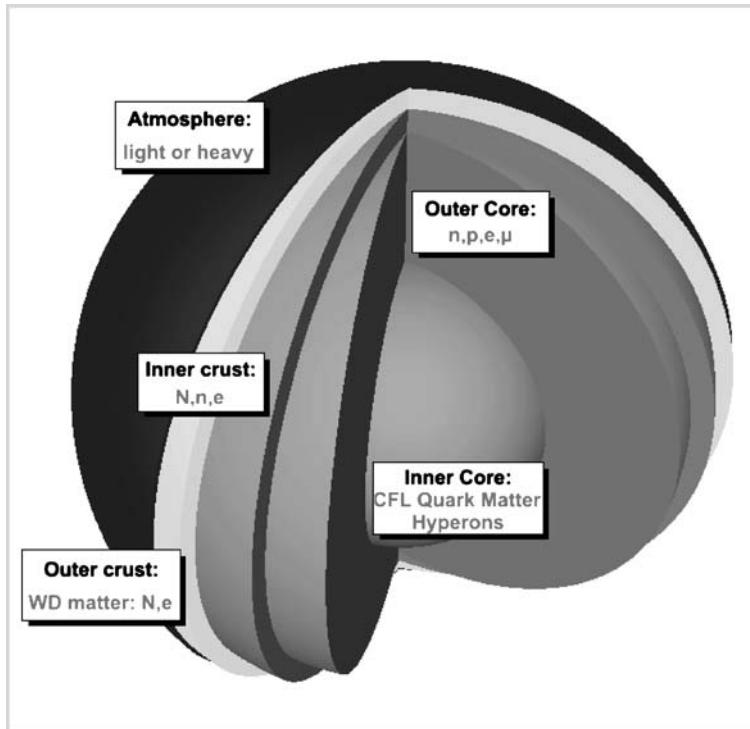


Fig. 6.1. Cross-section through the interior of a neutron star. The neutron star is surrounded by a thin atmosphere and an outer crust consisting of heavy nuclei and electrons. The inner crust consists of nuclei, neutrons and electrons, which at nuclear density make a transition to a neutron fluid. The composition of the central core is still unclear, but certainly consists in the outer part only of neutrons, protons, electrons and muons

where the core mainly consists of neutrons, protons and electrons. At high densities, however, also heavier baryons are excited, the neutron star now becomes a **hyperon star**. Since these baryons are so densely packed, a quark bag could be formed, and quarks are probably in a color-superconducting state. Finally, Bose condensates of pions and K mesons might occur. All these different internal structures lead to different mass–radii relations. For given mass, the traditional neutron star has the biggest radius, while neutron stars including quark cores are found to be more compact. Strange stars have the smallest radii.

6.2 Equations of State beyond Neutron Drip

The equation of state of dense matter plays an important role in the supernova phenomenon and in the structure and evolution of neutron stars. Matter in the collapsing core of massive stars at the end of its life is compressed from white

dwarf-like densities of about 10^6 g cm^{-3} to two or three times the nuclear saturation density, about $3 \times 10^{14} \text{ g cm}^{-3}$ (Lattimer and Prakash [249]). Nuclear saturation corresponds to a baryon density $n_s \simeq 0.16 \text{ baryons fm}^{-3}$. The central densities of neutron stars may even range up to $(3\text{--}10) n_s$. At densities around n_s and below, matter may be regarded as a mixture of neutrons, protons, electrons, positrons, neutrinos and antineutrinos, and photons. At higher densities, additional constituents may be present, such as muons, hyperons, kaons, pions and quarks. Up to now, there is no consensus on the composition of ultradense matter.

The main problem is to establish the state of the nucleons, which may be either bound in nuclei or be essentially free in continuum states. Neither temperature, nor densities are large enough to excite degrees of freedom, such as hyperons, mesons or quarks. Electrons are rather weakly interacting and may be treated as an ideal Fermi gas. At these high densities, the electrons are relativistic. As we have seen, at higher densities, the neutron chemical potential increases to the extent that the density of nucleons outside nuclei become large. These nucleons will modify the nuclear surface, decreasing the surface tension. At finite temperatures, even nuclear excited states become populated, and these states must be included by treating nuclei as *warm drops* of nuclear matter. At low temperatures, nucleons in nuclei are degenerate and the Fermi liquid theory is probably adequate for their description.

The fact that at subnuclear densities the spacing between nuclei may be of the same order of magnitude as the nuclear size itself will lead to substantial modifications. At zero temperature, the Wigner–Seitz description may be adequate for many applications. In this sense, one has to bridge the low-density and low-temperature regime, in which the nuclei can be described in terms of simple mass formula, with high densities and/or high temperatures in which the matter is a uniform bulk fluid. For this purpose, the *compressible liquid droplet model* for nuclei is useful, where the drop maintains thermal, mechanical and chemical equilibrium with its surroundings. Such a model was originally designed for finite temperature by Lattimer and Swesty [248].

6.2.1 From Neutron Drip to Saturation

If nuclear forces alone determined the equilibrium nuclear structure, nucleons would accumulate into nuclei of unlimited size. However, the Coulomb repulsive forces become so strong that these nuclei would undergo fission. We can therefore write the energy density of a mixture consisting of nuclei, free electrons and free neutrons in the form for the *compressible liquid drop model*

$$\epsilon = \epsilon(A, Z, n_N, n_n, n_e) = n_N(W_N + W_L) + \epsilon(n_e) + \epsilon_n(n_n)[1 - V_N n_N]. \quad (6.1)$$

In the compressible liquid drop model of nuclear matter, the energy of a nucleus is given by, x the proton fraction,

$$W_N = A \left[(1-x)m_n c^2 + xm_p c^2 + W(n, x) \right] + W_C + W_S. \quad (6.2)$$

W_C is the Coulomb energy and W_S the surface energy, while $W(n, x)$ denotes the bulk energy per nucleon depending on nucleon number density n . The quantity n_N denotes the number density of nuclei, while n_n is the number density of free neutrons. $\epsilon(n_e)$ is the electron energy. The quantity V_N is the volume of a nucleus, so that V_{NN} is the fraction of volume occupied by nuclei, and $1 - V_{NN}$ the fraction occupied by the neutron gas. The baryon density is then given by

$$n = An_N + n_n[1 - V_N n_N] \quad (6.3)$$

and the electron density by

$$n_e = Zn_N. \quad (6.4)$$

Nuclear matter represents *saturated systems*. The reason for this is in the short range of nuclear force, its strong repulsion at short distances and the Pauli principle. *Saturation means that, as more nucleons are added to nuclei, the density of the central region remains constant.* The nuclear radius therefore just scales as $R = r_n A^{1/3}$ with r_n as a constant. The energy density is therefore also constant. Accordingly, we can write the volume energy of a nucleus as the energy density of symmetric matter at normal density ϵ_n times the volume. This must be modified by the empirical fact that nuclei with neutron number N close to proton number Z are more tightly bound than their neighbors, altered by the increasing importance of the Coulomb repulsion at higher Z . The repulsion shifts the energy minimum of nuclei with increasing A to those with a greater neutron fraction $(A - Z)/Z = N/Z$.

In addition to these volume contributions, there will be a surface energy, which is also repulsive. This contribution arises because nucleons at the surface interact with fewer neighbors and so feel less attraction than those in the interior. This surface energy will be proportional to the surface area. Finally, there will also be a repulsive Coulomb energy because of the protons. As a consequence, the total mass of a nucleus can be written approximately as

$$M(A, Z) = A \left[\frac{4\pi}{3} r_n^3 \epsilon_n + a_{\text{sym}} \left(\frac{N - Z}{A} \right)^2 \right] + 4\pi r_n^2 A^{2/3} \epsilon_{\text{surf}} + \frac{3}{5} \frac{e^2 Z^2}{r_n A^{1/3}}. \quad (6.5)$$

This formula is sometimes referred to as the *semi-empirical mass formula*, or as the droplet model of nuclear masses. In actual practice, there are many more effects with unknown coefficients; these are incorporated in the best fit of parameters of a more general formula (Möller et al. [300]). As a result, the binding energy $B/A \simeq M(A, Z)/A - m_u$ is a series in various powers of $1/A$. With increasing A , just the volume term survives (which is constant). In this way, one discovers the binding energy per nucleon of infinite, symmetric nuclear matter and the symmetry coefficient

$$B/A = -16.3 \text{ MeV} , \quad a_{\text{sym}} = 32.5 \text{ MeV}. \quad (6.6)$$

The radius parameter $r_n = 1.16$ fm is obtained by an analysis of electron–nucleus scattering in terms of density distributions. The equilibrium value of the nucleon number density is related to r_n by

$$n_n = \left(\frac{4\pi}{3} r_n^3 \right)^{-1} = 0.153 \text{ fm}^{-3}. \quad (6.7)$$

In terms of an equation of state $\epsilon(n)$, the binding energy per nucleon is related to the equilibrium values of the saturation energy ϵ_n and the baryon number density n by

$$\frac{B}{A} = \left(\frac{\epsilon}{n} \right)_n - m \quad (6.8)$$

where $m = 938.93 \text{ MeV} = 4.7582 \text{ fm}^{-1}$ is the average of the neutron and proton masses. Thus, we get the energy density at saturation

$$\epsilon_n = 141 \text{ MeV fm}^{-3}. \quad (6.9)$$

The *compression modulus* K defined by

$$K = \left[P^2 \frac{d^2}{dP^2} \left(\frac{\epsilon(P)}{n(P)} \right) \right]_n = 9 \left[n^2 \frac{d^2}{dn^2} \left(\frac{\epsilon(n)}{n} \right) \right]_n \quad (6.10)$$

contains additional information on nuclear matter at saturation. The larger the value of K , the more steeply the equation of state will increase with density. The compression modulus likely lies in the range of 200 to 300 MeV, with recent evidence from a large body of data suggesting a value of $K = 234$ MeV.

The EoS by Baym, Pethick and Sutherland [64] improved Salpeter's calculation by using a somewhat improved mass formula and by also including the lattice energy (Fig. 6.2)

$$\epsilon = n_N M(A, Z) + \epsilon'(n_e) + \epsilon_n(n_n) + \epsilon_{\text{lattice}} \quad (6.11)$$

with

$$\epsilon_{\text{lattice}} = -1.444 Z^2 e^2 n_e^{4/3}. \quad (6.12)$$

From the energy density $\epsilon = \epsilon(n)$ one can derive the pressure

$$P = P(n) = n^2 \frac{\partial}{\partial n} \left(\frac{\epsilon}{n} \right). \quad (6.13)$$

A refined treatment was then given by Baym, Bethe and Pethick (BBP 1971 [65]) including results from detailed many-body calculations. These calculations end when nucleon–nucleon interaction become important, i.e. around nuclear densities.

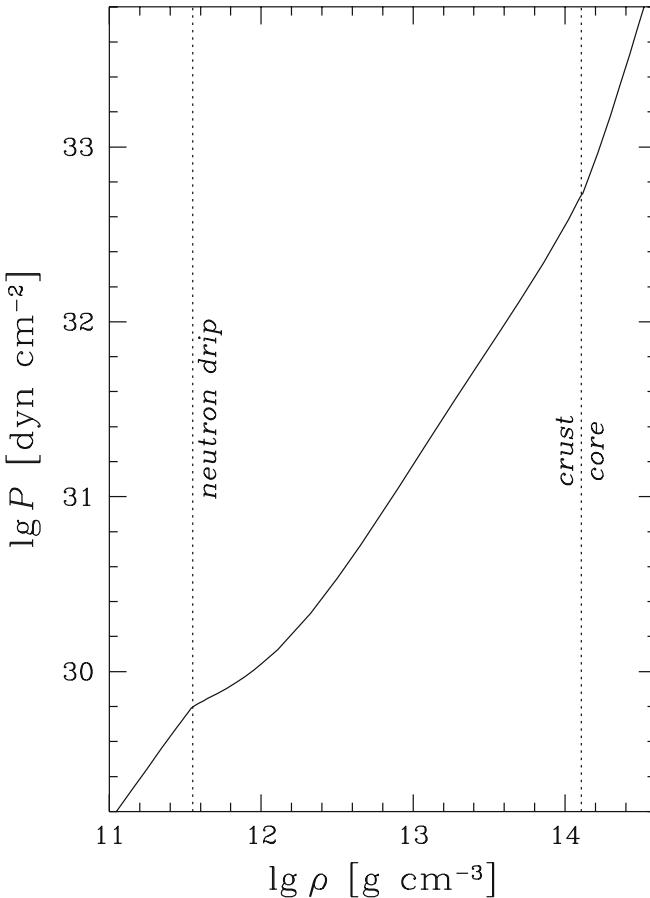


Fig. 6.2. Equation of state for cold nuclear matter [136]. Up to the neutron drip density, relativistic electrons are the main source for the pressure, $P \propto \varrho^{4/3}$, with some minor modifications. Beyond this density, the nonrelativistic neutrons overtake the pressure. In this region, strong forces are attractive, so that the effective pressure is somewhat below the pressure of the neutrons. Beyond the nuclear saturation density of $2 \times 10^{14} \text{ g cm}^{-3}$, nuclear forces become strongly repulsive

Empirical Bulk Energy of Nucleon Matter

Empirically, the energy per particle of nuclear matter reaches a minimum of about -16 MeV at a density $n_s = 0.16 \text{ fm}^{-3}$. Therefore close to this density its density dependence is roughly parabolic. The nucleon–nucleon interaction is optimized for symmetric matter (equal number of protons and neutrons, $x = 1/2$), so a parabolic dependence on the proton fraction x can be assumed. For analytical purposes, the nucleon free energy per baryon $W(n, x)$ can be approximated in MeV as

$$W(n, x) \simeq -16 + S_v(n)(1-2x)^2 + \frac{K_s}{18} \left(\frac{n}{n_s} - 1 \right)^2 - \frac{K'_s}{27} \left(\frac{n}{n_s} - 1 \right)^3 + F(T). \quad (6.14)$$

Here $F(T)$ includes finite temperature effects. S_v is the volume symmetry coefficient, numerically $S_v \simeq 30$ MeV, and K_s denotes the compressibility at nuclear density, $K_s = (200\text{--}300)$ MeV, with its most probable value $K_s = 234$ MeV. The experimental determination of these parameters comes from comparisons of total masses and energies of giant resonances of lab nuclei.

Nucleon–nucleon interactions can be included in the EoS by constructing a simple model for the nuclear potential that reproduces the general features of normal nuclear matter. The von-Weizäcker mass formula (6.5) for nuclides gives an equilibrium number density n_0 of 0.16 fm^{-3} for normal symmetric nuclear matter (i.e. $N = Z$). For this value of n_0 , the Fermi momentum $p_F^0 = 263 \text{ MeV}/c$ is small enough compared with $m_N = 939 \text{ MeV}/c^2$ so that a nonrelativistic treatment is sufficient. At this density, the average binding energy per nucleon $BE = -16 \text{ MeV}$ and the compressibility is between 200 and 300 MeV.

For this discussion, we introduce an empirical interaction for symmetric matter ($n_n = n_p$) of the following form

$$\frac{\epsilon(n)}{n} = m_N + \langle E_0 \rangle u^{2/3} + \frac{A}{2} u + \frac{B}{1+\sigma} u^\sigma, \quad u = n/n_0. \quad (6.15)$$

The quantities A and B , and the parameter σ are fit parameters. $\langle E_0 \rangle$ is the average kinetic energy per nucleon of symmetric matter in the ground state

$$\langle E_0 \rangle = \frac{3}{5} \frac{p_0^2}{2m_N} = \frac{3}{5} \frac{1}{2m_N} \left(\frac{3\pi^2 \hbar^3 n_0}{2} \right)^{2/3}. \quad (6.16)$$

From the three constraints, $d(\epsilon/n)/dn = 0$, the binding energy BE and the compressibility condition evaluated at nuclear density $u = 1$, we get three relations

$$\frac{2}{3} \langle E_0 \rangle + \frac{A}{2} + \frac{B\sigma}{1+\sigma} = 0 \quad (6.17)$$

$$\langle E_0 \rangle + \frac{A}{2} + \frac{B}{1+\sigma} = BE \quad (6.18)$$

$$\frac{10}{9} \langle E_0 \rangle + A + B\sigma = \frac{K_0}{9}. \quad (6.19)$$

These equations can be solved to yield

$$\sigma = \frac{K_0 + 2\langle E_0 \rangle}{3\langle E_0 \rangle - 9BE} \quad (6.20)$$

$$B = \frac{1+\sigma}{\sigma-1} \left(\frac{1}{3}\langle E_0 \rangle - BE \right) \quad (6.21)$$

$$A = BE - \frac{5}{3}\langle E_0 \rangle - B. \quad (6.22)$$

For $K_0 = 400$ MeV, one obtains the values with $\langle E_0 \rangle = 22.1$ MeV

$$A = -122.2 \text{ MeV}, \quad B = 65.39 \text{ MeV}, \quad \sigma = 2.112. \quad (6.23)$$

The corresponding pressure follows from the expression (6.13).

In neutron stars, we however find **nonsymmetric nuclear matter**. We represent the neutron and proton densities in terms of a parameter α defined as follows

$$n_n = \frac{1 + \alpha}{2} n, \quad n_p = \frac{1 - \alpha}{2} n. \quad (6.24)$$

This parameter means

$$\alpha = \frac{n_n - n_p}{n} = \frac{N - Z}{A}. \quad (6.25)$$

An alternative notation is

$$x = \frac{n_p}{n} = \frac{1 - \alpha}{2}. \quad (6.26)$$

For the kinetic energy one gets

$$\epsilon_{KE}(n, \alpha) = \frac{3}{5} \frac{p_{F,n}^2}{2m_N} n_n + \frac{3}{5} \frac{p_{F,p}^2}{2m_N} n_p = n \langle E_F \rangle \frac{1}{2} \left[(1 + \alpha)^{5/3} + (1 - \alpha)^{5/3} \right], \quad (6.27)$$

where

$$\langle E_F \rangle = \frac{3}{5} \frac{\hbar^2}{2m_N} \left(\frac{3\pi^2 n}{2} \right)^{2/3} \quad (6.28)$$

is the mean kinetic energy of symmetric nuclear matter at density n . For nonsymmetric matter, the excess kinetic energy is

$$\begin{aligned} \Delta\epsilon_{KE}(n, \alpha) &= \epsilon_{KE}(n, \alpha) - \epsilon_{KE}(n, 0) \\ &= n \langle E_F \rangle \left\{ \frac{1}{2} \left[(1 + \alpha)^{5/3} + (1 - \alpha)^{5/3} \right] - 1 \right\} \\ &= n \langle E_F \rangle \left\{ 2^{2/3} \left[(1 - x)^{5/3} + x^{5/3} \right] - 1 \right\}. \end{aligned} \quad (6.29)$$

For pure neutron matter, $\alpha = 1$, we find

$$\Delta\epsilon_{KE}(n, 1) = n \langle E_F \rangle (2^{2/3} - 1). \quad (6.30)$$

This gives the following Taylor expansion

$$\Delta\epsilon_{KE}(n, \alpha) = n \langle E_F \rangle \frac{5}{9} \alpha^2 \left(1 + \frac{\alpha^2}{27} + \dots \right). \quad (6.31)$$

For most considerations it is sufficient to keep terms to order α^2 . This gives a bulk symmetry energy of $\simeq 20$ MeV, and we may try an ansatz of the form

$$\frac{E(n, \alpha)}{A} = \frac{E(n, 0)}{A} + \alpha^2 S(n). \quad (6.32)$$

The isospin-symmetry breaking is proportional to α^2 , which roughly reflects the pairwise nature of the nucleon interaction. From the above, we may therefore assume the following form for $S(u)$

$$S(u) = (2^{2/3} - 1)\langle E_F \rangle [u^{2/3} - F(u)] + S_0 F(u). \quad (6.33)$$

Here, $S_0 = 30$ MeV is the bulk asymmetry energy which describes the energy difference between pure neutron matter and normal symmetric nuclear matter at ground-state density n_0 . The function $F(u)$ must satisfy $F(1) = 1$ (so that $S(u = 1) = S_0$) and $F(0) = 0$ (so that $S(u = 0) = 0$). Besides these constraints, the function $F(u)$ is not strongly constrained from experiments. We can make the simple ansatz $F(u) = u$. From this we obtain the following energy density for neutron matter

$$\begin{aligned} \frac{\epsilon_{\text{emp}}}{n} = m_n &+ \langle E_F \rangle u^{2/3} + \frac{A}{2} u + \frac{B}{1+\sigma} u^\sigma \\ &+ (2^{2/3} - 1)\langle E_F \rangle (u^{2/3} - u) + S_0 u. \end{aligned} \quad (6.34)$$

With the above parameters for symmetric nuclear matter, this gives the energy density parametrized in terms of the nuclear density $u = n/n_0$

$$\frac{\epsilon_{\text{emp}}}{n} = (939.6 + 35.1 u^{2/3} - 42.1 u + 21.0 u^{2.112}) \text{ MeV}. \quad (6.35)$$

The corresponding EoS is then also given by the expression (6.13) in parametrized form.

The empirical EoS is shown in Fig. 6.3. At densities below nuclear density, attractive forces soften the EoS, at densities higher than nuclear density, repulsive forces stiffen the EoS. This empirical model is certainly not justified for densities higher than 2–3 times the nuclear density. At high densities, the repulsive forces are probably too strong. At high densities, the sound speed

$$\left(\frac{c_S^2}{c^2}\right) = \frac{dP}{d\epsilon} = \frac{dP/dn}{d\epsilon/dn} \quad (6.36)$$

will violate causality, which requires $c_S^2 \leq c^2$.

Skyrme-Type Interaction

For comparison, we also show the properties of nuclear matter by applying the Hartree–Fock method on a phenomenological nucleon–nucleon interaction, which is parametrized in a Skyrme-type interaction

$$V(x, y) = \delta^3(x - y) \left(\frac{1}{6} t_3 n - t_0 \right), \quad (6.37)$$

where t_0 describes the attractive two-particle nuclear interaction, and $t_3 n$ the repulsive, density dependent many-body interaction which is dominant at high densities. For a homogeneous system, the total energy is given by

$$E_{\text{HF}} = \frac{3}{5} \frac{p_p^2}{2m_n} N_p + \frac{3}{5} \frac{p_n^2}{2m_n} N_n + \frac{1}{2} \left(\frac{1}{6} t_3 n - t_0 \right) \frac{N_p^2/2 + 2N_p N_n + N_n^2/2}{V}. \quad (6.38)$$

Similar to the above, the phenomenological parameters t_0 and t_3 are determined from the binding energy at nuclear density

$$t_0 = 1024.1 \text{ MeV fm}^3, \quad t_3 = 14600.8 \text{ MeV fm}^6. \quad (6.39)$$

This gives for pure neutron matter

$$E_{\text{HF}} = \frac{3}{5} \frac{p_n^2}{2m_n} N_n + \frac{1}{4} \left(\frac{1}{6} t_3 n - t_0 \right) n. \quad (6.40)$$

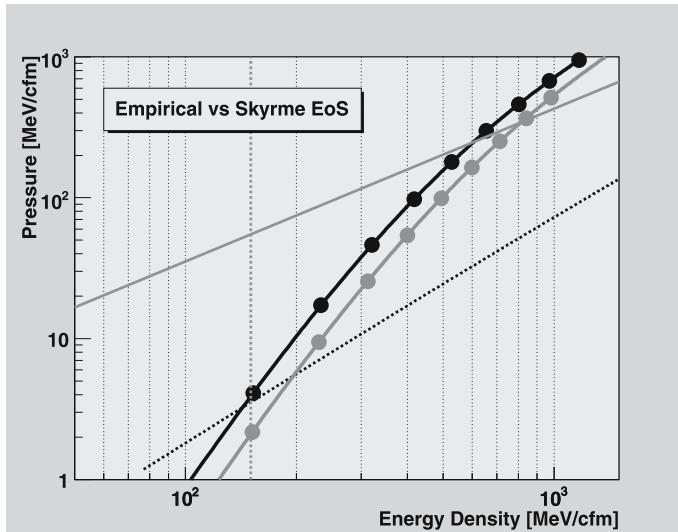


Fig. 6.3. EoS of nonsymmetric nuclear matter, energy density and pressure are given in units of MeV fm^{-3} . The *dashed curve* shows the free fermion gas for neutrons. The *solid line* corresponds to the empirical EoS. The *dots* denote the nucleon density in units of the saturated density n_0 , with values of $1.0, 1.5, 2.0, 2.5, \dots, 5 n_0$. The *upper hatched solid line* corresponds to the asymptotic free quark gas, $P = \epsilon/3$. The Skyrme model is represented by the *lower hatched solid line*

For the total energy density we then obtain

$$\epsilon(n) = m_n n + \frac{3}{10m_n} (3\pi^2 \hbar^3)^{2/3} n^{5/3} - \frac{t_0}{4} n^2 + \frac{t_3}{24} n^3 \quad (6.41)$$

and the corresponding expression for the pressure

$$P(n) = \frac{2}{10m_n} (3\pi^2 \hbar^3)^{2/3} n^{5/3} - \frac{t_0}{4} n^2 + \frac{t_3}{12} n^3. \quad (6.42)$$

The resulting EoS is plotted in Fig. 6.3. It is only slightly shifted with respect to the above empirical relation, but also violates causality at high densities.

The Crust and Nuclear Pasta

The EoS of the outer crust is rather well established. As soon as one leaves the region of experimentally known nuclei, the EoS of cold catalyzed matter becomes uncertain. This uncertainty rises at densities higher than the neutron drip density. The properties of nuclei are affected by the ambient neutron gas which contributes more and more to the total pressure. Therefore, the problem of correct modelling of the EoS of pure neutron gas at subnuclear densities becomes important. The real EoS of cold catalyzed matter stems from the real nucleon Hamiltonian, which is expected to describe nucleon interactions at $\varrho < 2\varrho_0$. In practice, in order to make the solution of the many-body problem feasible, the task is reduced to finding an *effective nucleon Hamiltonian*.

The properties of the inner crust EoS are illustrated in Fig. 6.2, using the results obtained by Douchin and Haensel [136] for the SLy model of effective nuclear hamiltonian; it will be hereafter referred to as the SLy4 EoS. Notice that the sound velocity is $v_s = \sqrt{dP/d\varrho}$. A significant softening (decrease of v_s) occurs just after the neutron drip point. At densities greater than the neutron drip one, the SLy4 EoS stiffens gradually (v_s increases) with growing density, due to the increasing contribution of dripped neutrons to the pressure. There is a discontinuous increase (jump) of sound velocity at the crust–core interface.

In the SLy4 EoS, the crust–core transition takes place as a very weak first-order phase transition, with the relative density jump of the order of a percent [136]. For this model, the spherical nuclei persist to the crust bottom. In contrary, for the FPS EoS, the crust–core transition takes place through a sequence of phase transitions with the changes of nuclear shapes (from spheres to spaghetti and lasagne, and finally to bubbles [259]). All in all, while the presence of the exotic nuclear shapes is expected to have dramatic effect on the transport phenomena and elastic properties of nuclear matter, their effect on the EoS is negligible.

At densities of about 0.05 fm^{-3} in the inner crust, nucleons are arranged in a variety of complex shapes, called nuclear pasta. Baryonic matter is organized as a result of short-range nuclear attraction and long-range Coulomb repulsion. Competition among these interactions plays a fundamental role in the organization of matter and results in a so-called Coulomb frustration. Frustration develops from

the inability of a system to simultaneously satisfy all of its elementary interactions. For example, the Ising antiferromagnet on a triangular lattice is frustrated, because not all of the nearest neighbor spins can be antiparallel to each other. At subnuclear densities of about 10^{14} g cm $^{-3}$, Coulomb frustration is expected to promote the development of complex shapes. These shapes follow from the competition between surface tension and Coulomb energies. Surface tension would favor spherical shapes, Coulomb interactions often favor nonspherical configurations. Therefore, a variety of complex structures – such as spheres, cylinders, plates, etc. – have been predicted. The many phases of nuclear matter displaying this variety of shapes are now known as **nuclear pasta**. The properties of this matter is of relevance to the structure of the inner crust of neutron stars and to the dynamics of core-collapse supernovae. The ground-state shapes of nuclear pasta have been calculated by various authors [273, 413, 414].

6.2.2 Nuclear EoS for Dense Neutron Matter

Modern models for the equation of state for the star's core fall into two categories: nonrelativistic variational approximation and relativistic field theoretic approaches. The first approach leads in general to acausal behavior at high densities, due to its basis on the nonrelativistic Schrödinger equation. The most realistic EoS based on nonrelativistic potentials are given in the following Table 6.1 (only neutrons, protons, electrons and muons, but no meson condensate and no hyperons).

At $\varrho \simeq \varrho_0$, core matter is a liquid composed mostly of neutrons with a few percent admixture of the equal number of protons and electrons. If the Fermi energy

Table 6.1. Equations of state of the liquid core of neutron star based on nonrelativistic modelling

EoS	Composition and model	Reference
BPAL12	$npe\mu$, effective nucleon energy functional	Bombaci et al. 1995 [82]
BGN1H1	$n p \Sigma \Lambda \Xi e \mu$, effective baryon energy functional	Balberg et al. [47]
BBB1	$npe\mu$, Brueckner theory, Argonne NN plus Urbana NNN potentials	Baldo et al. [52]
FPS	$npe\mu$, effective nucleon energy functional	Pandharipande and Ravenhall [323]
BGN2H1	$n p \Sigma \Lambda \Xi e \mu$, effective baryon energy functional	Balberg et al. [47]
BBB2	$npe\mu$, Brueckner theory, Paris NN plus Urbana NNN potentials	Baldo et al. [52]
SLy	$npe\mu$, effective nucleon energy functional	Douchin and Haensel [136]
BGN1	$npe\mu$, effective baryon energy functional	Balberg et al. [47]
APR	$npe\mu$, variational theory, Nijmegen NN plus Urbana NNN potentials	Akmal et al. [26]
BGN2	$npe\mu$, effective nucleon energy functional	Balberg et al. [47]

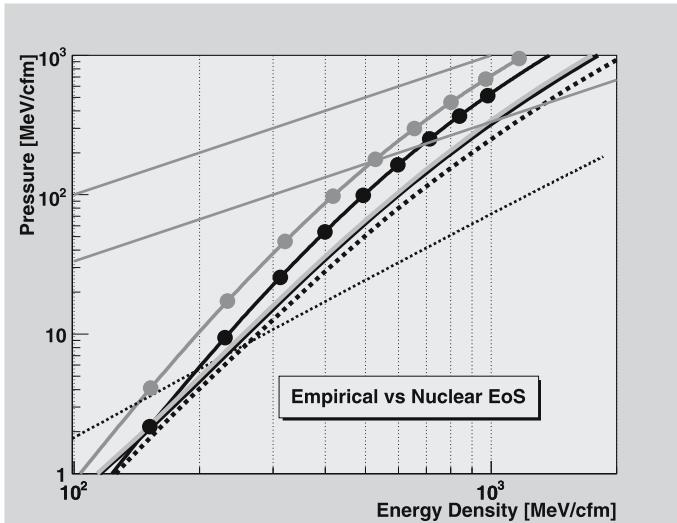


Fig. 6.4. Nuclear EoS FPS (thick dotted line) and SLy (solid hatched line) compared to the empirical and simple Skyrme models (with dots denoting the number density in units of nuclear density as in Fig. 6.3). The straight lines correspond to the stiffest EoS $P = \rho$ and the asymptotically free quark gas $P = \rho/3$

of electrons exceeds the muon rest energy (105.7 MeV), muons replace a fraction of electrons to minimize the energy of the system. Such a system in beta equilibrium is usually called $npe\mu$ matter. This is the simplest model of matter in neutron star cores: except for the presence of muons, which are insignificant for the EoS, the matter constituents – neutrons, protons and electrons – are the same as in familiar terrestrial matter. Still, even for this simplest composition, the uncertainties in the EoS are quite large, especially at densities significantly higher than ρ_0 . This results from the approximations and deficiencies of the many-body theory of dense nucleon matter, and from the lack of knowledge of strong interactions in superdense matter. These uncertainties are illustrated in Fig. 6.5. The brief characteristics of these models and the references to the original papers are given in Table 6.1. The BPAL12 and BGN2 EoSs should be considered as the soft and stiff extremes of the theoretical models. Notice that the stiffest BGN2 EoS is superluminal (sound velocity $v_s > c$) at the highest densities relevant for neutron stars, which reflects the inadequacy of the nonrelativistic approach at such a high density.

Configurations of hydrostatic equilibrium of nonrotating neutron stars will be calculated by solving the Tolman–Oppenheimer–Volkoff (TOV) for various EoS for nuclear matter (FPS, SLy, and APR). These models of matter at supranuclear densities are the most elaborate ones, and are based on experimental nuclear physics and relatively precise many-body calculations of dense neutron matter. They do not include possible dense matter constituents, for which strong interactions are poorly known (hyperons), or which are hypothetical (pion and kaon condensates, or quark

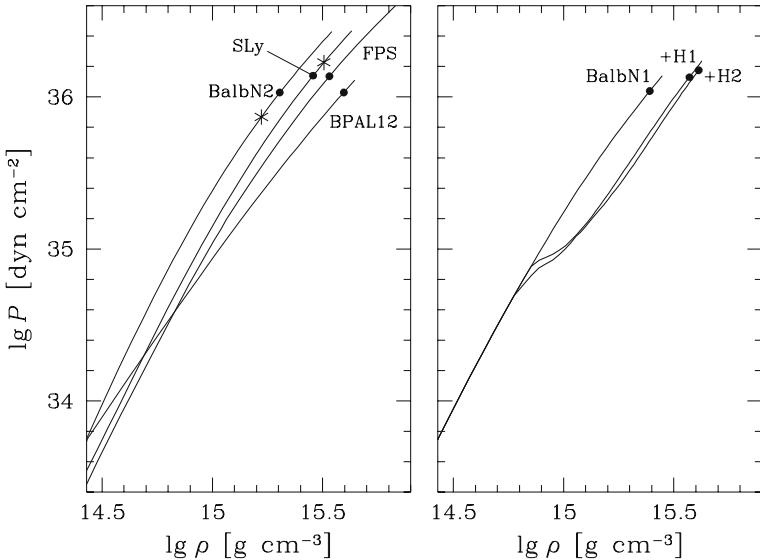


Fig. 6.5. Selected models for EoS of nonrelativistic nuclear matter in the core of a neutron star. BGN1 and BGN2 here have *labels* BalbN1 and BalbN2. Large dots correspond to maximum density in stable neutron stars. Asterisks correspond to the density above which EoS is superluminal ($v_s > c$). *Left panel:* EoSs of the $npe\mu$ matter. *Right panel:* effect of hyperons is shown by comparing the EoS without hyperons (i.e. for the $npe\mu$ matter) and EoSs in which hyperons Λ , Σ , Ξ are included (+H1 and +H2 correspond to the BGN1H1 and BGN1H2 models of Table 6.1). Figure adapted from [187]

matter). The TOV equations are integrated from the center of the configuration, with boundary condition at $r = 0$: $P(0) = P_c$, $m(0) = 0$. The stellar surface at $r = R$ is then determined by the pressure condition $P(R) = 0$. The total gravitational mass M is then given by $M = m(R)$, and the total number of baryons by $N = n(R)$.

Mass, Central Density, and Radius of Neutron Stars

Classical neutron stars based on EoS of FPS and SLy4 have rather shallow density profiles (Fig. 6.6 and Fig. 6.7). The core extends to about 10 km, before the density steeply falls off.

In Fig. 6.8 we show the dependence of gravitational mass on the central density, ρ_c , for $\rho_c > 3 \times 10^{14}$ g cm $^{-3}$, and compare it with that obtained for the softer EoS FPS [15]. Actually, on the lower-density side the curve exhibits a minimum at $M_{\min} \simeq 0.09 M_\odot$, not shown in the figure. The value of M_{\min} depends rather weakly on the EoS. On the higher-density side, $M(\rho_c)$ has a maximum. The existence of this maximum (for any EoS) is an important consequence of general relativity. Configurations with bigger central densities cannot exist in hydrostatic equilibrium and collapse into black holes. For SLy4, we get $M_{\max} \simeq 2.05 M_\odot$, to be compared with $M_{\max} = 1.80 M_\odot$ for a softer FPS EoS.

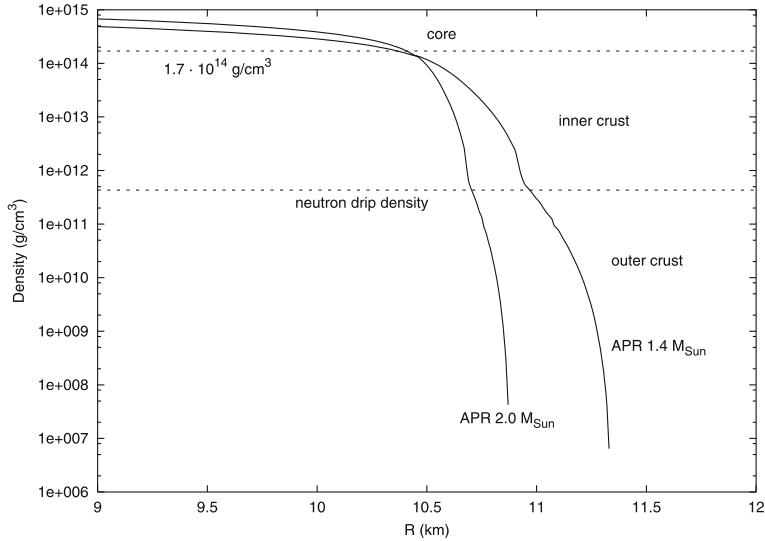


Fig. 6.6. Density profiles for classical neutron stars within the APR EoS. The various shells are clearly visible in the profile. Figure provided by A. Bauswein [63]

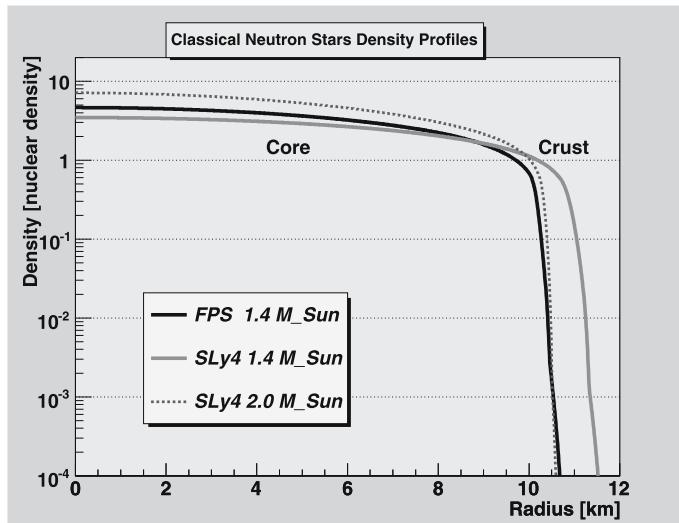


Fig. 6.7. Density profiles for classical neutron stars (for EoS FPS and SLy4, and different masses as indicated in the legend). Figure based on data provided by A. Bauswein [63]

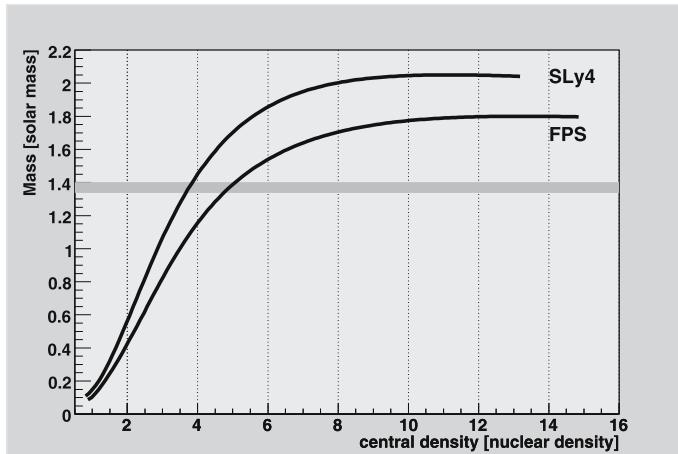


Fig. 6.8. Mass as a function of central density for classical EoS (FPS and SLy4). The *hatched* region corresponds to the mean mass found for neutron stars from binary pulsars. Figure based on data provided by A. Bauswein [63]

The central density of the maximum allowable mass configuration is the maximum one which can be reached within static neutron stars. Models with $\varrho_c > \varrho_c(M_{\max})$ have $dM/d\varrho_c < 0$. They are therefore unstable with respect to small radial perturbations and collapse into black holes. The maximum central density for static stable neutron stars is, for the nuclear EoS, $2.9 \times 10^{15} \text{ g cm}^{-3}$. The corresponding maximum value of baryon density is $n_{\max} = 1.21 \text{ fm}^{-3} \simeq 7.6 n_s$, to be compared with what is obtained for the FPS EoS.

A comparison with the APR98 EoS is also of interest, the mass curve obtained for the nuclear EoS is quite close to the APR98 one, especially for $1.4 M_\odot$. It should be mentioned that, for the APR EoS, neutron star models contain a central core with $v_{\text{sound}} > c$, which is unphysical. Such a problem does not arise for the nuclear EoS discussed here, for which $v_{\text{sound}} < c$ within all stable neutron star models.

Radius vs. Gravitational Mass

The radius–mass relation, obtained for EoS FPS and SLy4 for static, cold neutron stars, is shown in Fig. 6.9, where for the sake of comparison we also show the $R(M)$ curve for the FPS EoS. For masses between $1.0 M_\odot$ and $2.0 M_\odot$, the neutron star radius decreases rather weakly with increasing mass, from 12 km to 10 km. For neutron star masses, measured for some binary radio pulsars (shaded band), the radius would be slightly below 12 km. The insensitivity of R to M for $1 \leq M \leq 2.0 M_\odot$ is typical of the realistic EoS without a strong softening at high density.

In this figure, we also plot the radii as a function of mass for some measured surface redshift $z = 1/\sqrt{1 - 2GM/Rc^2} - 1$. The redshift measured by Cottam et al. [126] would require in fact a too high mass for the neutron star in EXO 0748–676. This is already an indication that neutron stars with masses above 1.3 solar

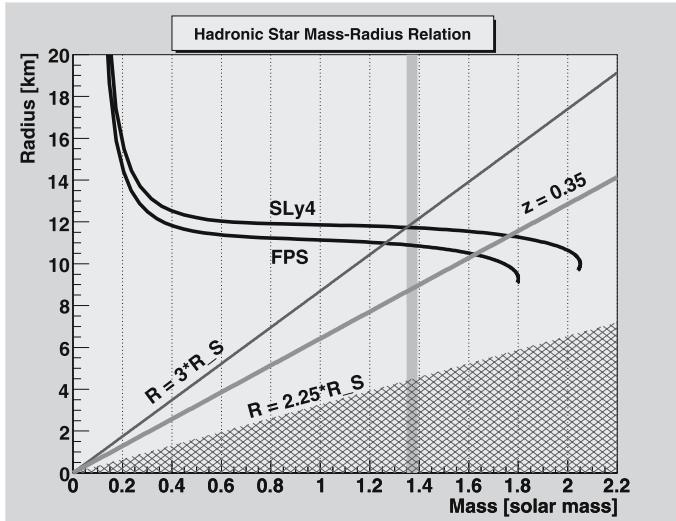


Fig. 6.9. Radii as a function of mass for classical EoS (FPS vs. SLy4). Neutron stars with masses above $1.35 M_{\odot}$ have radii smaller than the innermost stable orbit (ISCO). Figure based on data provided by A. Bauswein (LSW Heidelberg)

masses are probably not classical neutron stars. They require a softer EoS to satisfy all these requirements, unless the mass in LMXBs is indeed higher than observed for millisecond neutron stars. This is, however, implausible, since LMXBs are the progenitors of recycled millisecond pulsars.

For a static neutron star, general relativity predicts that the circular Keplerian orbits (for test particles) with $r > 3R_S$ are stable, and those with $r < 3R_S$ are unstable, where the gravitational radius $R_S = 2GM/c^2 = 2.95 M/M_{\odot}$ km. The radius of the marginally stable orbit, which separates these two classes of orbits, is therefore $r_{\text{ms}} = 3R_S = 12.4(M/1.4 M_{\odot})$ km. As we see in Fig. 6.9, for $M > 1.35 M_{\odot}$ we find $r_{\text{ms}} > R$ for classical neutron stars, and therefore for such neutron stars the innermost stable circular orbit (ISCO) is separated from the stellar surface by a gap. A similar situation also holds for the FPS EoS. Note that the existence of a gap between the ISCO and neutron star surface might be important for the interpretation of the spectra of the kilohertz quasiperiodic oscillations (QPOs) observed in the X-ray radiation of some low-mass X-ray binaries (see Sect. 6.6).

Surface Redshift

The surface redshift of photons emitted from neutron star photosphere is given by $z_{\text{surf}} = 1/\sqrt{1 - R_S/R} - 1$. At given M , the SLy value of this value is systematically lower than for softer FPS EoS (Fig. 6.10). However, the maximum surface redshift for the SLy EoS, 0.59, is some 10% higher than for the FPS EoS, 0.53. The larger value of the maximum mass for stiffer SLy EoS plays a decisive role in determining

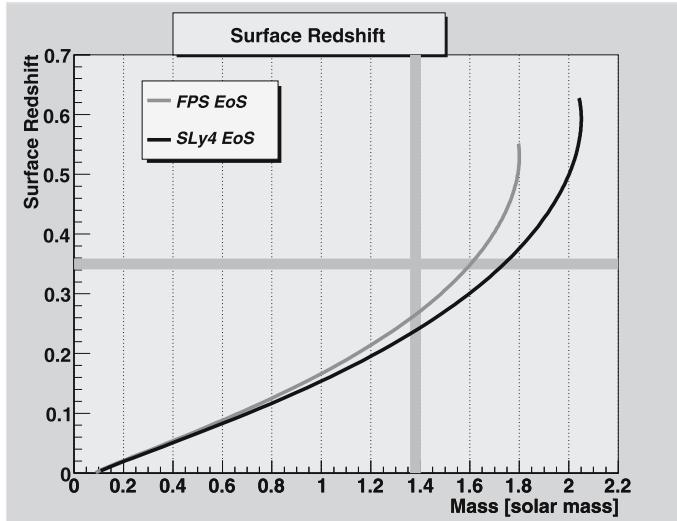


Fig. 6.10. Surface redshift as a function of the mass for classical EoS (FPS and SLy4). The hatched lines denote the mean mass measured by binary pulsars and the redshift for EXO 0748–676. Figure based on data provided by A. Bauswein (LSW Heidelberg)

the spacetime curvature close to neutron star with maximum allowable mass. In the range of measured values of masses of binary pulsars we get $z_{\text{surf}} \simeq 0.22\text{--}0.26$, slightly higher than for the FPS EoS. For $M < 2.0 M_{\odot}$, the curve for our EoS is quite similar to the APR98 one.

Binding Energy

The binding energy of a neutron star, E_{bind} , is defined as the mass defect with respect to a dispersed configuration of matter consisting of the same number of baryons, multiplied by c^2 . A dispersed configuration is characterized by negligible pressure and negligible gravitational interactions. Equivalently, one may define E_{bind} as a net work, needed to transform a neutron star into a dispersed configuration of matter. We use the standard definition of E_{bind} , i.e. with respect to a dispersed configuration of a pressureless cloud of ^{56}Fe dust, with mass per nucleon $m_{\text{Fe}} = 1.6587 \times 10^{-24}$ g. Therefore,

$$E_{\text{bind}} = (Nm_{\text{Fe}} - M)c^2. \quad (6.43)$$

With such a definition, E_{bind} represents a good approximation of the binding energy of neutron stars with respect to the configuration of a presupernova core from which the neutron star was formed, via gravitational collapse, as a by-product of the type II supernova explosion. At given M , it is somewhat smaller than for a softer FPS EoS. However, the maximum value of E_{bind} , reached for M_{max} , is significantly larger for the SLy EoS than for a softer FPS one. Binding of the neutron star is due to gravitational forces and it rises rapidly with M .

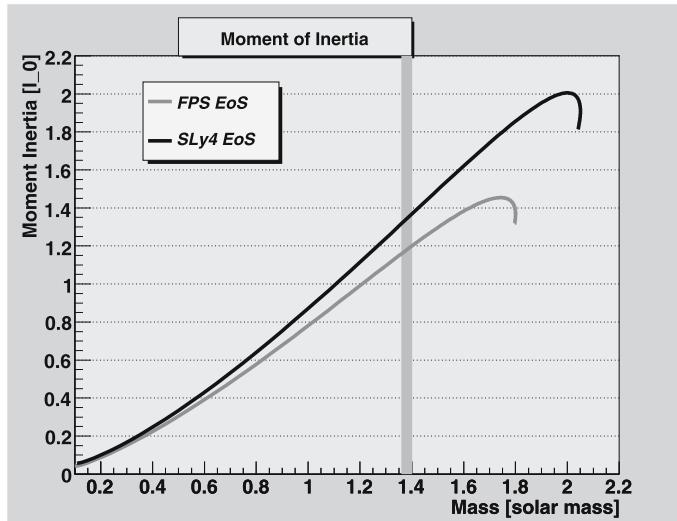


Fig. 6.11. Moment of inertia as a function of mass for classical EoS (FPS and SLy4), normalized to $I_0 = 1.0 \times 10^{45}$ cgs units. Figure based on data provided by A. Bauswein (LSW Heidelberg)

Moment of Inertia

Most observed neutron stars are rotating. However, even for the very rapid millisecond pulsar PSR 1937+21, with rotation period $P = 1.558$ ms, rotation implies only small changes of stellar structure for neutron stars with $M > 1.0 M_\odot$. Therefore, for the description of effects of rotation for observed neutron stars one can use a slow rotation approximation, in which effects of rotation (assumed to be rigid) are treated using a lowest order perturbative scheme (see Sect. 7.4). In this approach, one calculates, in the linear approximation in the angular frequency as measured by a distant observer, Ω , the total angular momentum of a neutron star, $J \propto \Omega$. Then, one gets the moment of inertia for slow, rigid rotation as $I = J/\Omega$. Notice that within the slow rotation approximation, I is independent of Ω and can be calculated from the structure of a nonrotating configuration of a neutron star. The values of I are plotted, vs. M , in Fig. 6.11. At given M , the value of I for the nuclear EoS is significantly higher than for softer FPS EoS. The difference rises rapidly with increasing M .

6.2.3 Relativistic Mean Field Theory above Saturation

At nuclear densities and beyond, the nuclei overlap so that the entire central core of a neutron star forms a kind of giant nucleus. The free neutrons dominate the equation of state. In hadronic matter at densities of several times that of normal nuclear matter, the nucleon Fermi energy will be sufficiently large that it is energetically more favorable to some nucleons to transform to heavier baryon species through the

Table 6.2. Baryon octet and mesonic states relevant for neutron star matter

Particle	m (MeV)	Spin	Isospin	B	Q	S
N	939	1/2	1/2	1	0,1	0
Λ	1115	1/2	0	1	0	-1
Σ	1190	1/2	1	1	-1,0,1	-1
Ξ	1315	1/2	1/2	1	-1,0	-2
σ	800	0	0	0	0	0
ω	782	1	0	0	0	0
ϱ	770	1	1	0	-1,0,1	0
π	139	0	1	0	-1,0,1	0
K^+	494	0	1/2	0	1	1
K^0	494	0	1/2	0	0	1
\bar{K}^0	494	0	1/2	0	0	-1
K^-	494	0	1/2	0	-1	-1

electroweak interaction. In this way, the conserved baryon number is shared among a greater number of species of lower Fermi energies.

Nucleons have a radius of about 0.8 fm (compare with the charge radius determined by scattering experiments on protons). If such nucleons would be packed into a cubic lattice to the point of touching, the density would be 0.24 fm^{-3} , or 1.5 times nuclear density. These nucleons are surrounded by a cloud of mesons with an extension of about double size of the charge radius, i.e. about 2 fm. At a few times nuclear density, these mesonic clouds overlap and they have to be included in the interaction between the nucleons. These mesons are built by the vector particles ϱ and ω , and the scalar particle σ . In this description, hadronic matter can be considered as a kind of fluid consisting of various types of hadrons, leptons and mesons. At higher densities, even the hard cores of the nucleons can overlap so that hadronic matter undergoes a phase transition to the quark–gluon plasma. The exact value for the density of this transition is still unknown.

The most important baryons in neutron stars are the octet baryons which are the lowest mass baryons of spin 1/2. They comprise the nucleons and some of the hyperons. The most important mesons are the scalar σ , the spin-one vector ω and the spin and isospin vector–isovector ϱ . In a mean-field approximation, the above mesons mediate the interaction among baryons. The pion and kaon are of possible interest as condensates. Since baryons and mesons are not themselves the fundamental particles they were once thought to be, at higher densities the quark constituents come into play, and QCD is the appropriate theory to deal with. The hadron–quark transition density in low-temperature matter is, however, still under debate, values are between 2 and 8 times the nuclear density.

A treatment in terms of a relativistic field theory has some merits (see the discussion in Glendenning [5]). First of all, it is automatically causal, whereas the EoS computed on the basis of Schrödinger’s equation is not compelled to be.

Secondly, the coupling constants of relativistic field theory can easily be related to the bulk properties of nuclear matter. And finally, at high densities, the asymptotic freedom of quark matter is naturally built in.

To describe the complex baryonic composition of neutron star matter, the Lagrangian of hadronic matter has to be sufficiently general (in this section we use the signature convention $(+ - -)$ of high-energy physics and also the natural units $\hbar = 1 = c$, for conversion use $\hbar c = 197.33 \text{ MeV fm}$)

$$\begin{aligned} \mathcal{L} = & \sum_B \bar{\Psi}_B [i\gamma^\mu \partial_\mu - m_B + g_{\sigma B} \sigma - g_{\omega B} \gamma^\mu \omega_\mu - \frac{1}{2} g_{\varrho B} \gamma^\mu \mathbf{q}_\mu \cdot \boldsymbol{\tau}] \Psi_B \\ & + \frac{1}{2} \partial^\mu \sigma \partial_\mu \sigma - \frac{1}{2} m_\sigma^2 \sigma^2 - \frac{\kappa m_\sigma}{3} (g_\sigma \sigma)^3 - \frac{\lambda}{4} (g_\sigma \sigma)^4 \\ & - \frac{1}{4} Q_{\mu\nu} Q^{\mu\nu} + \frac{1}{2} m_\omega^2 \omega^\mu \omega_\mu + \frac{\zeta}{4} g_\omega^4 (\omega^\mu \omega_\mu)^2 \\ & - \frac{1}{4} \mathbf{R}_{\mu\nu} \cdot \mathbf{R}^{\mu\nu} + \frac{1}{2} m_\varrho^2 \mathbf{q}^\mu \cdot \mathbf{q}_\mu + \frac{\xi}{4} g_\varrho^4 (\mathbf{q}^\mu \cdot \mathbf{q}_\mu)^2 \\ & + \sum_{L=e,\mu} \bar{\Psi}_L (i\gamma^\mu \partial_\mu - m_L) \Psi_L . \end{aligned} \quad (6.44)$$

The baryon spinors are denoted by Ψ_B with the adjoint $\bar{\Psi}_B = \Psi^+ \gamma^0$, γ^μ are the Dirac matrices and $\boldsymbol{\tau}$ the Pauli matrices¹. In this model for hadronic matter, the baryons interact via the exchange of σ , ω and ϱ mesons, including also hyperons (Müller and Serot [5, 305, 410]). The first line gives the sum of the baryon Lagrangians and their interactions with the scalar, vector and vector-isovector mesons. The sum is taken over all the charge states of the baryon octet ($p, n, \Lambda, \Sigma^+, \Sigma^0, \Sigma^-, \Xi^-, \Xi^0$). The second line contains the Lagrangian of the scalar meson including self-interaction. Their interactions with the baryons give rise to the attraction. The third line describes the vector meson which is responsible for the repulsive force between baryons, given by the massive vector field ω_μ with its field tensor

$$Q_{\mu\nu} = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu . \quad (6.45)$$

The fourth line gives the Lagrangian for the isovector particle which couples to the isospin of the baryons and is responsible for the isospin symmetry energy, $\mathbf{q}_\mu = (q_\mu^1, q_\mu^2, q_\mu^3)$. Its field tensor is given by

$$R_{\mu\nu}^a = \partial_\mu q_\nu^a - \partial_\nu q_\mu^a - g_\varrho \epsilon^{abc} q_\mu^b q_\nu^c . \quad (6.46)$$

The last line contains the Lagrangian for the leptons (electrons and muons) which are required for the charge neutrality of neutron star matter. The masses have the dimension 1/fm, the vector fields also and $\bar{\Psi}\Psi$ has the dimension of a particle density².

¹ Pauli matrices: $\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

² $1/\text{fm}^4 = 197.33 \text{ MeV/fm}^3 = 3.5178 \times 10^{14} \text{ g cm}^{-3} = 3.1616 \times 10^{35} \text{ dyne/cm}^2$.

One can regard this Lagrangian as an effective one in which the quantum effects of the vacuum fluctuations have already been included. For this reason, there are a number of parameters which describe the effective coupling.

With this effective Lagrangian one can calculate the density effects in nuclear matter. Generally, one uses the **mean field approach**. In uniform nuclear matter, the ground-state expectation value of the spatial components of the ω meson will vanish. We only have $\langle \omega^0 \rangle$ as expectation value, which is usually denoted by ω . Similarly, for the ϱ meson field, only the time-like neutral component $\langle \varrho_0^3 \rangle$ does not vanish. Also, the expectation value of the σ field is denoted by $\sigma \neq 0$, in general. A nonvanishing expectation value in σ leads to a reduction of the nucleon mass m_B , $m^* = m_B - g_\sigma \sigma$.

This Lagrangian density is a function of fields $\Phi(x)$ and their derivatives $\partial_\mu \Phi$, $\mathcal{L} = \mathcal{L}[\Phi(x), \partial_\mu \Phi(x)]$. The relativistic action is therefore dimensionless

$$S = \int_{t_0}^{t_1} d^4x \mathcal{L}[\Phi(x), \partial_\mu \Phi(x)]. \quad (6.47)$$

The variation of this action produces then the Euler–Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \Phi(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi(x))} = 0. \quad (6.48)$$

Since the Lagrangian depends on several fields, there is one such equation for each field. In distinction to the Schrödinger equation, this formalism guarantees the covariance, and therefore the Lorentz invariance of the theory.

In particular, the invariance of this action against translational symmetry

$$x'^\mu = x^\mu + \epsilon^\mu \quad (6.49)$$

leads to conservation of the energy–momentum tensor

$T^{\mu\nu} \equiv \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi(x))} \partial^\nu \Phi - \eta^{\mu\nu} \mathcal{L}$

(6.50)

implied by the field equations

$$\partial_\mu T^{\mu\nu} = 0, \quad \nu = 0, 1, 2, 3. \quad (6.51)$$

The expectation values of the energy–momentum tensor provide then the energy density and the pressure of the system.

The σ – ω Model

In order to illustrate the procedure for calculating the EoS we consider first a much simpler model. The interaction by means of scalar particles and vector particles (so-called σ – ω model) goes back to Johnson and Teller (1955), Duerr (1956) and Walecka

(1974). This model is based on four particles, two nucleons, $\Psi = (\Psi_p, \Psi_n)$, a scalar meson (for the Yukawa potential) and the omega vector meson for the repulsive force (self-interaction is neglected)

$$\begin{aligned} \mathcal{L} = & \bar{\Psi} [i\gamma^\mu(\partial_\mu + ig_\omega\omega_\mu) - (m_N - g_\sigma\sigma)] \Psi \\ & + \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma - \frac{1}{2}m_\sigma^2\sigma^2 - \frac{1}{4}\omega_{\mu\nu}\omega^{\mu\nu} + \frac{1}{2}m_\omega^2\omega^\mu\omega_\mu. \end{aligned} \quad (6.52)$$

The Euler–Lagrange equations for the mesons are

$$(\square + m_\sigma^2)\sigma(x) = g_\sigma \bar{\Psi}\Psi(x) \quad (6.53)$$

and for the vector meson

$$(\square + m_\omega^2)\omega_\mu(x) - \partial_\mu\partial^\nu\omega_\nu(x) = g_\omega \bar{\Psi}\gamma_\mu\Psi. \quad (6.54)$$

In the mean field approximation, the values of the mesonic fields are given in terms of their expectation values with the result

$$m_\sigma^2\langle\sigma\rangle = g_\sigma\langle\bar{\Psi}\Psi\rangle \quad (6.55)$$

$$m_\omega^2\langle\omega_\mu\rangle = g_\omega\langle\bar{\Psi}\gamma_\mu\Psi\rangle. \quad (6.56)$$

The nucleon fields themselves are momentum eigenstates, $\Psi(x) = \Psi(k) \exp(-ik \cdot x)$. The Dirac equation leads then to the equation

$$[\gamma_\mu(k^\mu - g_\omega\omega^\mu) - (m - g_\sigma\langle\sigma\rangle)]\Psi(k) = 0. \quad (6.57)$$

With the definition of $K^\mu \equiv k^\mu - g_\omega\langle\omega^\mu\rangle$, this leads to the eigenvalue equation

$$(K^\mu K_\mu - m^{*2})\Psi(K) = 0. \quad (6.58)$$

Here we introduced the **effective mass** $m^* \equiv m - g_\sigma\langle\sigma\rangle$. From this we can calculate the four-vector $k = (k^0, \mathbf{k})$ by

$$e(\mathbf{k}) \equiv k^0 = K^0 + g_\omega\langle\omega_0\rangle. \quad (6.59)$$

Hence the nucleon eigenvalues of the three-momentum \mathbf{k} for particles and antiparticles are

$$e(\mathbf{k}) = E(\mathbf{k}) + g_\omega\langle\omega_0\rangle \quad (6.60)$$

$$\bar{e}(\mathbf{k}) = E(\mathbf{k}) - g_\omega\langle\omega_0\rangle, \quad (6.61)$$

with

$$E(\mathbf{k}) = K_0 = \sqrt{(\mathbf{k} - g_\omega\omega_0)^2 + (m - g_\sigma\sigma)^2}. \quad (6.62)$$

We have found now the Dirac momentum eigenvalues in terms of the meson fields. But these meson fields are also given in terms of the Dirac charges, this poses

therefore a self-consistent problem. Let us denote the expectation value of an operator in the single-particle state as follows $(\bar{\Psi} \Gamma \Psi)_k$. Then the expectation value of this operator in the ground state of the many-nucleon system is given by

$$\langle \bar{\Psi} \Gamma \Psi \rangle = \sum \int \frac{d^3 k}{(2\pi)^3} (\bar{\Psi} \Gamma \Psi)_k \Theta(\mu - e(\mathbf{k})). \quad (6.63)$$

The sum is understood as summation over all spin and isospin states of the occupied momentum states. $\Theta(x)$ is the step-function, and we denote the Fermi energy by μ (also called the chemical potential). The ground-state expectation values are momentum integrals over all filled states with eigenvalues $e(\mathbf{k})$ below the chemical potential. The fermion surface is described by the relation $e(\mathbf{k}) = \mu$.

There is a tricky method to calculate the expectation values (Glendenning [5]). The Hamiltonian corresponding to the Dirac equation is

$$H_D = \gamma_0 \left[\gamma \cdot \mathbf{k} + g_\omega \gamma^\mu \omega_\mu + m^* \right]. \quad (6.64)$$

The expectation value in a single state is therefore

$$(\Psi^+ H_D \Psi)_k = K_0(\mathbf{k}) = E(\mathbf{k}) + g_\omega \omega_0. \quad (6.65)$$

Now we take a derivative of the left-hand side with respect to any variable ζ in the Hamiltonian

$$\frac{\partial}{\partial \zeta} (\Psi^+ H_D \Psi)_k = \left(\Psi^+ \frac{\partial H_D}{\partial \zeta} \Psi \right)_k \quad (6.66)$$

Since $\Psi(\mathbf{k})$ is an eigenfunction, it does not depend on ζ . As an example we take the derivative with respect to ω_0 which yields the normalization

$$(\Psi^+ \Psi)_k = 1 \quad (6.67)$$

and therefore, we find for the baryon density

$$n_B = \langle \Psi^+ \Psi \rangle = 4 \int_0^{k_F} \frac{d^3 k}{(2\pi)^3} \Theta(\mu - e(\mathbf{k})) = \frac{2k_F^3}{3\pi^2}. \quad (6.68)$$

This relates the Fermi momentum to the baryon density. In the same manner one can take the derivative with respect to k^i and can calculate other expectation values

$$(\bar{\Psi} \gamma^i \Psi)_k = \frac{\partial}{\partial k^i} E(\mathbf{k}) \quad (6.69)$$

and hence

$$\langle \bar{\Psi} \gamma^i \Psi \rangle = 4 \int_0^{k_F} \frac{d^3 k}{(2\pi)^3} \frac{\partial E(\mathbf{k})}{\partial k^i} \Theta(\mu - e(\mathbf{k})) = 0. \quad (6.70)$$

For the source term of the sigma meson we also need the scalar density $\bar{\Psi}\Psi$. Its expectation value follows from the derivative with respect to m

$$(\bar{\Psi}\Psi)_k = \frac{\partial E}{\partial m} \quad (6.71)$$

The result is now

$$n_s = \langle \bar{\Psi}\Psi \rangle = \frac{2}{\pi^2} \int_0^{k_F} k^2 dk \frac{m - g_\sigma \langle \sigma \rangle}{\sqrt{k^2 + (m - g_\sigma \langle \sigma \rangle)^2}}. \quad (6.72)$$

Combining these results, we find the implicit equations for the expectation values σ and ω^μ (in the following we identify the fields with their expectation values, $\sigma \equiv \langle \sigma \rangle$)

$$g_\sigma \sigma = \left(\frac{g_\sigma}{m_\sigma} \right)^2 \frac{2}{\pi^2} \int_0^{k_F} k^2 dk \frac{m - g_\sigma \sigma}{\sqrt{k^2 + (m - g_\sigma \sigma)^2}} \quad (6.73)$$

$$g_\omega \omega_0 = \left(\frac{g_\omega}{m_\omega} \right)^2 n_B \quad (6.74)$$

$$\omega_i = 0. \quad (6.75)$$

The first equation is a transcendental equation for σ . In fact one can show that the low-density and high-density limits can be approximated by means of

$$g_\sigma \sigma \simeq m \left(\frac{g_\sigma k}{\pi m_\sigma} \right)^2 \left[1 + \left(\frac{g_\sigma k}{\pi m_\sigma} \right)^2 \right]^{-1}. \quad (6.76)$$

Therefore, the effective mass $m^* = m - g_\sigma \sigma$ has the vacuum value m at low densities and it tends to zero at high densities.

To compute the equation of state in this model we use the expressions for the energy-momentum tensor whose expectation value in the rest frame of matter is diagonal

$$\langle T^{\mu\nu} \rangle = \text{diag}(\varrho, P, P, P). \quad (6.77)$$

These expectation values can be expressed in terms of the Lagrangian

$$\varrho = -\langle \mathcal{L} \rangle + \langle \bar{\Psi} \gamma_0 k_0 \Psi \rangle \quad (6.78)$$

$$P = \langle \mathcal{L} \rangle + \frac{1}{3} \langle \bar{\Psi} \gamma_i k_i \Psi \rangle. \quad (6.79)$$

The expectation values of the mesonic fields are independent of spacetime. The expectation value for the Lagrangian is simply given by

$$\langle \mathcal{L} \rangle = -\frac{1}{2} m_\sigma^2 \sigma^2 - \frac{1}{2} m_\omega^2 \omega_0^2. \quad (6.80)$$

From the above results we also get

$$\begin{aligned}\langle \bar{\Psi} \gamma_0 k_0 \Psi \rangle &= \frac{2}{\pi^2} \int_0^{k_F} k^2 dk e(\mathbf{k}) \\ &= \frac{2}{\pi^2} \int_0^{k_F} k^2 dk \left[g_\omega \omega_0 + \sqrt{k^2 + (m - g_\sigma \sigma)^2} \right] \\ &= m_\omega^2 \omega_0^2 + \frac{2}{\pi^2} \int_0^{k_F} k^2 dk \sqrt{k^2 + (m - g_\sigma \sigma)^2}.\end{aligned}\quad (6.81)$$

Similarly, we have

$$(\bar{\Psi} \gamma \Psi)_k = \frac{\partial E(\mathbf{k})}{\partial \mathbf{k}} \quad (6.82)$$

and therefore

$$\langle \bar{\Psi} \gamma \cdot \mathbf{k} \Psi \rangle = \frac{2}{\pi^2} \int_0^{k_F} \frac{k^4}{\sqrt{k^2 + (m - g_\sigma \sigma)^2}} dk. \quad (6.83)$$

This yields for the pressure

$$P = -\frac{1}{2} m_\sigma^2 \sigma^2 + \frac{1}{2} m_\omega^2 \omega_0^2 + \frac{1}{3} \frac{2}{\pi^2} \int_0^{k_F} \frac{k^4}{\sqrt{k^2 + (m - g_\sigma \sigma)^2}} dk.$$

(6.84)

and for the energy density

$$\varrho = \frac{1}{2} m_\sigma^2 \sigma^2 + \frac{1}{2} m_\omega^2 \omega_0^2 + \frac{2}{\pi^2} \int_0^{k_F} k^2 dk \sqrt{k^2 + (m - g_\sigma \sigma)^2}.$$

(6.85)

The first terms are contributions from the mesonic fields and the integrals give the contributions from the filled up nucleon states. For $\sigma = 0 = \omega_0$, we recover the standard expressions for a free fermion gas (see, e.g. the expressions for the free electron gas in the last section). The expression for the pressure shows explicitly that **scalar mesons produce a softening in the pressure** (negative contribution), while the **vector meson contributes a hardening** (positive contribution). In this way we obtain the equation of state as $\varrho(n)$ and $P(n)$. n can be numerically eliminated in order to obtain $P(\varrho)$. Because $n_s \propto k^3$, the vector terms dominate at high density with $P \rightarrow \varrho$ from below. As a consequence, the speed of sound $\sqrt{dP/d\varrho} \rightarrow 1$ from below, i.e. this theory is causal. In addition, the coupling parameters in this theory are then related to the observational parameters of nuclear matter.

The Isospin Current

The isospin of the baryons produces an additional interaction. The above theory has to be extended to incorporate an isospin restoring interaction, an effect which is

visible in the mass formula. We need to introduce a meson that has as its source the three-component of the isospin. For this purpose one can use the rho meson which appears as an isospin triplet field with the isospin current

$$\mathbf{I}^\nu = \boldsymbol{\varrho}_\mu \times \boldsymbol{\varrho}^{\nu\mu}. \quad (6.86)$$

The isospin current of the nucleon fields is then

$$\mathbf{I}^\mu = \frac{1}{2} \bar{\Psi} \gamma^\mu \boldsymbol{\tau} \Psi \quad (6.87)$$

with the isospin charge

$$I_3^0 = \frac{1}{2} \bar{\Psi} \gamma^0 \tau_3 \Psi = \frac{1}{2} (\Psi_p^+ \Psi_p - \Psi_n^+ \Psi_n) = \frac{1}{2} (n_p - n_n). \quad (6.88)$$

This charge density is the source for the mean mesonic field

$$g_\varrho \langle \varrho_3^0 \rangle = \left(\frac{g_\varrho}{m_\varrho} \right)^2 \frac{1}{2} (n_p - n_n) \quad (6.89)$$

$$\langle \varrho_3^k \rangle = 0. \quad (6.90)$$

So the asymmetry is now responsible for a nonvanishing rho field.

A Complete Hadronic Model

When the self-interaction in the scalar field is included, everything goes a bit more complicated. But the integrals can still be done numerically (see Glendenning [5]). In the solutions of the full Euler–Lagrange equations the mesonic fields are replaced by their mean values in static uniform matter, and the nucleon currents are expressed in terms of their ground-state expectation values generated in the presence of the mean meson fields. This procedure is called the **mean field approximation** of nuclear field theory. The coupling constants are related to the parameters of nuclear matter at saturation (Table 6.3). In this sense, the equation for the fermion field of each species can be given in the momentum space

$$\left[\gamma^\mu \left(k_\mu - g_{\omega B} \omega_\mu - \frac{1}{2} g_{\varrho B} \tau_3 \varrho_{3,\mu} \right) - (m_B - g_{\sigma B} \sigma) \right] \Psi_B(k) = 0. \quad (6.91)$$

The eigenvalues of particle and antiparticle are similarly given

$$e_B(\mathbf{k}) = g_{\omega B} \omega_0 + g_{\varrho B} \varrho_{03} I_{3B} + \sqrt{k^2 + (m_B - g_{\sigma B} \sigma)^2} \quad (6.92)$$

$$\bar{e}_B(\mathbf{k}) = -g_{\omega B} \omega_0 - g_{\varrho B} \varrho_{03} \bar{I}_{3B} + \sqrt{k^2 + (m_B - g_{\sigma B} \sigma)^2}. \quad (6.93)$$

I_{3B} is the isospin three-component of each baryon. Here we have used that only the three-component of the isospin field survives in the mean-field approximation. Similarly, the meson field equations are

Table 6.3. The five coupling constants that yield the correct binding energy $B/A = -16.3$ MeV, the saturation density $n_s = 0.153 \text{ fm}^{-3}$ and the symmetry energy $a_{\text{asym}} = 32.5 \text{ MeV}$ as a function of the compressibility and the effective mass m^* (Glendenning [5])

K MeV	m^*/m	$(g_\sigma/m_\sigma)^2$ fm^2	$(g_\omega/m_\omega)^2$ fm^2	$(g_\varrho/m_\varrho)^2$ fm^2	κ $\times 100$	λ $\times 100$
200	0.70	12.684	7.148	4.410	0.5610	-0.6986
200	0.75	11.299	5.696	4.656	0.8784	-1.0098
200	0.80	9.926	4.233	4.876	1.4602	-1.2412
250	0.70	12.230	7.148	4.410	0.4312	-0.4103
250	0.75	10.727	5.696	4.656	0.6275	-0.3409
250	0.80	9.134	4.233	4.876	0.8804	0.6917
300	0.70	11.785	7.148	4.410	0.2948	-0.1071
300	0.75	10.177	5.696	4.656	0.3601	0.3722
300	0.80	8.403	4.233	4.876	0.2480	2.7997

$$\omega_0 = \sum_B \frac{g_{\omega B}}{m_\omega^2} n_B \quad (6.94)$$

$$\varrho_{03} = \sum_B \frac{g_{\varrho B}}{m_\varrho^2} I_{3B} n_B \quad (6.95)$$

$$m_\sigma^2 \sigma = -\kappa m_\sigma g_\sigma (g_\sigma \sigma)^2 - \lambda g_\sigma (g_\sigma \sigma)^3 + \sum_B \frac{2J_B + 1}{2\pi^2} g_{\sigma B} \int_0^{k_B} \frac{m_B - g_{\sigma B} \sigma}{\sqrt{k^2 + (m_B - g_{\sigma B} \sigma)^2}} k^2 dk. \quad (6.96)$$

The energy density and the pressure now follow in the same manner as in the $\sigma-\omega$ model from the expectation value of the energy-momentum tensor

$$\varrho = \frac{1}{2} m_\sigma^2 \sigma^2 + \frac{1}{2} m_\omega^2 \omega_0^2 + \frac{1}{2} m_\varrho^2 \varrho_{03}^2 + \frac{1}{3} \kappa m_\sigma (g_\sigma \sigma)^3 + \frac{1}{4} \lambda (g_\sigma \sigma)^4 + \sum_B \frac{2J_B + 1}{2\pi^2} \int_0^{k_B} \sqrt{k^2 + (m_B - g_{\sigma B} \sigma)^2} k^2 dk \quad (6.97)$$

$$P = -\frac{1}{2} m_\sigma^2 \sigma^2 + \frac{1}{2} m_\omega^2 \omega_0^2 + \frac{1}{2} m_\varrho^2 \varrho_{03}^2 - \frac{1}{3} \kappa m_\sigma (g_\sigma \sigma)^3 - \frac{1}{4} \lambda (g_\sigma \sigma)^4 + \frac{1}{3} \sum_B \frac{2J_B + 1}{2\pi^2} \int_0^{k_B} \frac{k^4 dk}{\sqrt{k^2 + (m_B - g_{\sigma B} \sigma)^2}}. \quad (6.98)$$

The nonlinearities in the sigma field, given by the coupling constants κ and λ , and the nonvanishing expectation value of the rho meson also contribute to the equation of state. Here, we have neglected the self-interaction terms for the vector fields. The number density of each species follows from

$$n_B = \frac{2J_B + 1}{6\pi^2} b_B k_B^3 \quad (6.99)$$

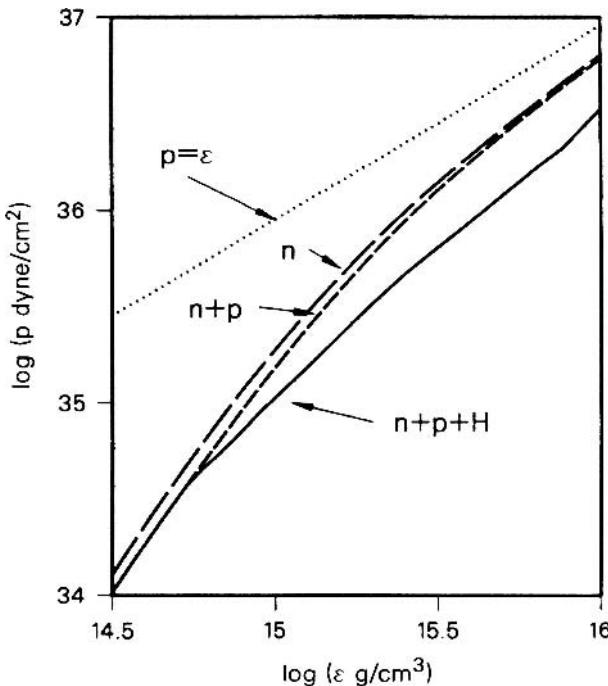


Fig. 6.12. EoS for pure neutron matter (n), neutrons in equilibrium with protons and electrons ($n+p$), and in equilibrium with hyperons ($n+p+H$). Here $K = 240$ MeV and $m^* = 0.78$ m. Figure adapted from [5]

for spin J_B and the baryon charge b_B . The total density is then $n = \sum_B n_B$. The above two equations determine the equation of state in parametrized form. The solutions of these nonlinear relations result in the Fig. 6.12 for various values of the compressibility and the effective mass for the most probable EoS).

In any theory having a vector meson interaction one finds for large baryon density the asymptotic behavior $P \rightarrow \varrho$. As in the $\sigma-\omega$ model, this can explicitly be shown for the complete hadronic model. The pressure approaches the energy density from below, as required by causality arguments – the speed of sound is always less than the speed of light in such theories. In fact, the correct behavior will be even softer, since at high densities hadronic matter makes a transition to quark matter, where $P \rightarrow \varrho/3$ in the high-density limit.

The low-density limit of this theory is of no interest (i.e. below saturation density). In this limit, the ground state of hadronic matter is no longer a uniform gas of nucleons. As discussed above, it is more favorable for nucleons to clump together in the form of nuclei and to form a lattice of nuclei filled up with free neutrons and relativistic electrons. By lowering the baryon density, the system undergoes a phase transition.

6.2.4 Analytical Fits to EoS

For numerical purposes, it is sometimes useful to have suitable analytical fits to EoS. Haensel and Potkin [188] have recently given such a scheme. A particular

EoS is usually presented in the form of a table containing a grid of calculated matter constituents divided by c^2 , $\varrho = \epsilon/c^2$, baryon number density n and pressure P . The EoS $\{\varrho_i, n_i, P_i\}$, $i = 1, \dots, N$, is then interpolated between the tabulated forms to get a one-parameter form $P = P(n)$ and $\varrho = \varrho(n)$. Depending on the interpolation schemes, different results for the mass-radius relation are in general obtained. Since the first law of thermodynamics requires

$$P = nc^2 \frac{d}{dn} \left(\frac{\varrho}{n} \right) \quad (6.100)$$

we may invert this relation to get

$$\frac{\varrho}{n} = \frac{\varrho_*}{n_*} + \int_{n_*}^n \frac{P(n')}{n'c^2} dn', \quad (6.101)$$

where n_* is the value of n at the neutron star surface. This is set equal to the density of ^{56}Fe at zero pressure and zero temperature, the corresponding value $\varrho_* = 7.6 \text{ g cm}^{-3}$. In the outermost neutron star layers, we fix the mass per nucleon, $m_0 = 1.66 \times 10^{-24} \text{ g}$, so that $n_* = \varrho_*/m_0 = 4.73494 \times 10^{24} \text{ cm}^{-3}$.

As we have seen, there are three qualitatively different interior regions in a neutron star, separated by phase transition points: the outer crust (consisting of nuclei and electrons), the inner crust (consisting of electrons, nuclei and dripped neutrons), and the core which contains neutrons, electrons and protons, μ^- mesons, and possibly π and K condensates, some hyperons, or even quark matter. Haensel and Potekhin [188] rely their fit on the tabulated EoS FPS (or SLy) below $\varrho > 5 \times 10^{10} \text{ g cm}^{-3}$. At extremely low densities, the EoS would depend on temperature, this can however be neglected for fitting. They propose the following fit with $\xi = \log(\varrho[\text{g cm}^{-3}])$ and $\zeta = \log(P[\text{dyn cm}^{-2}])$

$$\begin{aligned} \zeta = & \frac{a_0 + a_2\xi + a_3\xi^3}{1 + a_4\xi} F_0[a_5(\xi - a_6)] \\ & +(a_7 + a_8\xi) F_0[a_9(a_{10} - \xi)] \\ & +(a_{11} + a_{12}\xi) F_0[a_{13}(a_{14} - \xi)] \\ & +(a_{15} + a_{16}\xi) F_0[a_{17}(a_{18} - \xi)]. \end{aligned} \quad (6.102)$$

The parameters a_i are given in Table 6.4. The typical fit error is 1%–2%. The function F_0 is the Fermi function

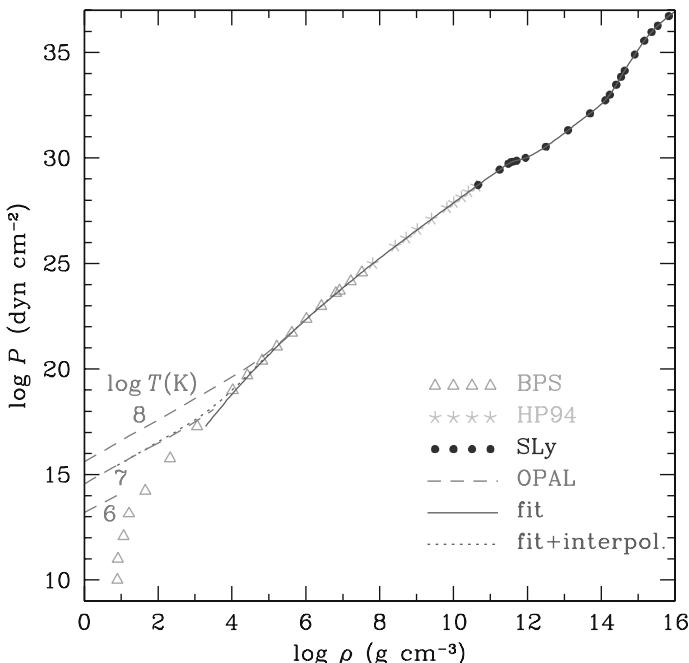
$$F_0(x) = \frac{1}{\exp(x) + 1}. \quad (6.103)$$

The comparison of the fit and the data is presented in Fig. 6.13. The typical fit error of ϱ is about 1% at $\eta > 10^{-7}$ (corresponding to $\xi > 3$), and the maximum error < 4% occurs near the neutron drip and crust–core phase transitions.

It is also instructive to treat ϱ and P as a function of baryon number density n . For this one can use the following fit

Table 6.4. Analytic nuclear EoS [188]

i	$a_{i,FPS}$	$a_{i,SLy}$	i	$a_{i,FPS}$	$a_{i,SLy}$
1	6.22	6.22	10	11.8421	11.4950
2	6.121	6.121	11	-22.003	-22.775
3	0.006004	0.005925	12	1.5552	1.5707
4	0.16345	16.326	13	9.3	4.3
5	6.50	6.48	14	14.19	14.08
6	11.8440	11.4971	15	23.37	27.80
7	17.24	19.105	16	-1.508	-1.653
8	1.065	0.8938	17	1.79	1.50
9	6.54	6.54	18	15.13	14.67

**Fig. 6.13.** Analytic fits to neutron star EoS for nonrotating configurations. In the outer crust, the EoS is modified by finite temperature effects. Figure adapted from Haensel and Potekhin [188]

$$\frac{\varrho}{nm_0} = 1 + \frac{p_1 n^{p_2} + p_3 n^{p_4}}{(1 + p_5 n)^2} F_0[-p_6(\log n + p_7)] + \frac{n}{8 \times 10^{-6} + 2.1 n^{0.585}} F_0[p_6(\log n + p_7)]. \quad (6.104)$$

n is given in units of baryons fm $^{-3}$.

In Fig. 6.14 we plot the energy density as a function of n . For $n < 0.1 \text{ fm}^{-3}$, ϱ is proportional to n . Only in the core, the total energy density exceeds considerably the rest-mass energy density. In the outer crust, the value of Γ depends quite weakly on density. At $10^8 \leq \varrho \leq 3 \times 10^{11} \text{ g cm}^{-3}$, this value would be $\Gamma \approx 4/3$ if A, Z values were fixed, because in this case P is mainly determined by the pressure of ultrarelativistic electron gas which behaves as $\propto (Z\varrho/A)^{4/3}$. For example, $\Gamma \approx 4/3$ within each shell with constant A and Z . However, the compressible liquid drop model used by [136] effectively smoothes the discontinuities caused by transitions from one to another (A, Z) species with increasing density, which leads to an effective continuous increase of the A/Z ratio and corresponding decrease of Γ , seen in Fig. 6.14.

A dramatic drop in Γ occurs at neutron drip threshold, which corresponds to strong softening of the EoS. The analytical expression somewhat smoothes this drop. The behavior of Γ in the inner crust results from an interplay of several factors, with stiffening due to interaction between dripped neutrons, a softening effect of the coexistence of neutron gas and nuclear matter, and the softening Coulomb contribution.

At the crust–core interface, matter strongly stiffens, and Γ jumps from ≈ 1.7 to ≈ 2.2 , which results from disappearance of nuclei. The analytical approximation also smoothes this jump, though reflects the stiffening. This approximation is also smooth across a small discontinuous drop of Γ at $\varrho \approx 2 \times 10^{14} \text{ g cm}^{-3}$, where muons start to replace a part of ultrarelativistic electrons. However, the electrons and muons give only minor contribution to the pressure (and therefore behavior of Γ) in the core, because the main contribution comes from interactions between nucleons.

6.3 Neutron Star Models

Now having such equations of state at hand for high densities, we can study the corresponding solutions of the TOV equations. These solutions are called neutron stars, though the interior does not entirely consist of neutron matter. Neutron stars are bound by gravity, not by nuclear forces. Gravity is the binding force for large objects. This binding energy per nucleon due to gravity is of the order of 160 MeV/A compared to the binding energy of nuclear matter at its saturation density $\simeq 16 \text{ MeV/A}$. In addition, gravity compresses the matter of neutron stars to densities above the nuclear density. At these high densities, the nucleons feel only the repulsive force.

6.3.1 Hadronic Models

Chemical Composition of Neutron Stars

Neutron stars are born hot, at temperatures in the range of 20 MeV. Neutrinos and photons produced in the interior diffuse however to the surface, leave the star, lowering thereby its energy content. In the hot state, also hyperons and mesons are produced by weak interactions. The temperature falls however in a short time below

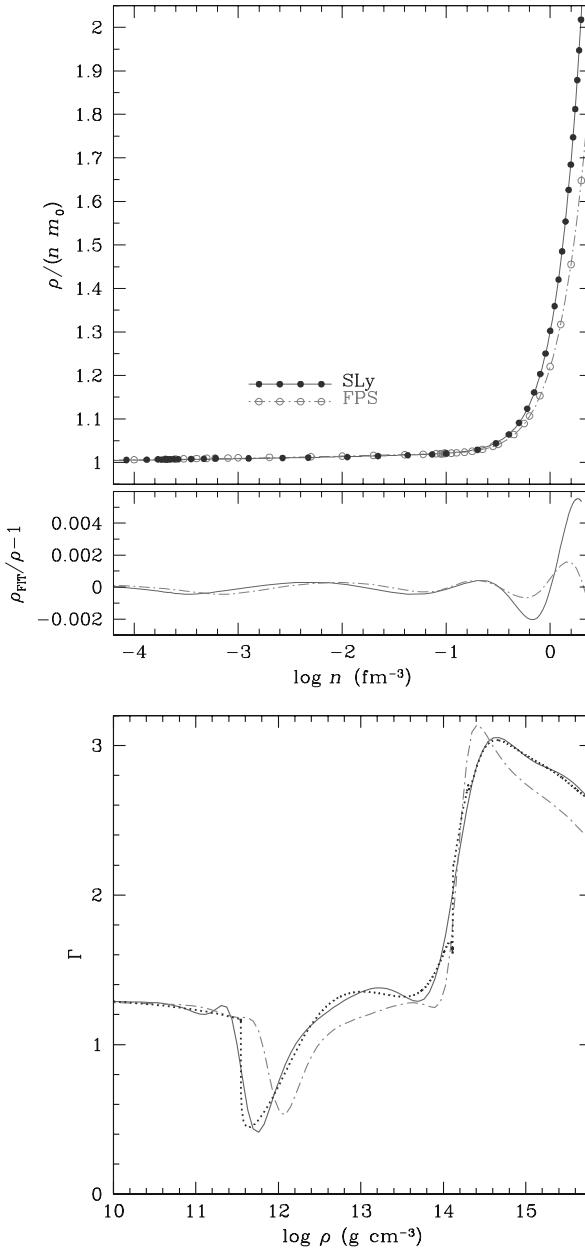


Fig. 6.14. Dependence of ϱ and Γ as a function of n for the analytic neutron star EoS [188]. *Upper panel:* Dependence of ϱ as a function of n for the analytic neutron star EoS (rarefied data (symbols) and the fit (lines)), with the relative difference between the data and fit. *Lower panel:* Adiabatic index Γ for SLy and FPS EoSs. *Solid line:* analytical approximation (SLy); *dotted line:* precise values (SLy); *dot-dashed line:* analytical approximation (FPS). Figures adapted from Haensel and Potekhin [188]

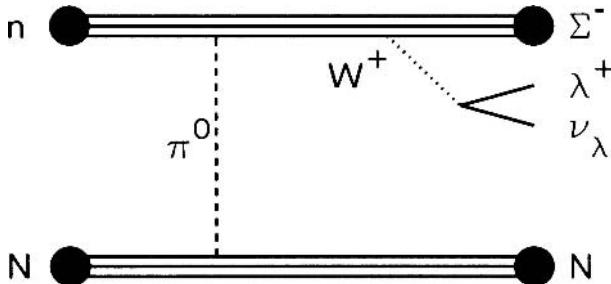


Fig. 6.15. The weak flavor changing reaction at the quark level

10^9 K. In the hot phase, the fermions are not degenerate so that strong interactions might occur

$$N + N \rightarrow N + \Lambda + K . \quad (6.105)$$

Some of these particles decay over various channels

$$K^0 \rightarrow 2\gamma \quad (6.106)$$

$$K^- \rightarrow \mu^- + \bar{\nu} \quad (6.107)$$

$$\mu^- + K^+ \rightarrow \mu^- + \mu^+ + \nu \rightarrow 2\gamma + \nu . \quad (6.108)$$

As the temperature falls, these reactions are no longer possible. However, the growth of strangeness still continues through direct weak flavor changing reactions such as shown in Fig. 6.15

$$n + N + \mu^- \rightarrow N + \Sigma^- + 2\gamma + \nu . \quad (6.109)$$

A particle type is only populated, if its chemical potential exceeds its lowest energy state in the medium. In lowest approximation, this would be its vacuum mass. The masses of the particles are however affected by their interactions with other particles in the medium, this in general lowers the mass of the particle. From this we expect that superdense matter will be populated by many baryonic species, possibly even with quarks.

There could also exist a **condensate** of charged negative mesons. In particular, K^- condensates have been discussed in the literature as a possible high-density phase of nuclear matter. Being a boson, all kaons could populate the same momentum state. The condensation by π^- may be another possibility.

These particles can then no longer decay, since Pauli blocking occurs. For temperatures below 10^9 K, the star is cold and has reached its ground state. Because neutrinos are lost from the star, their chemical potential vanishes, $\mu_\nu = 0$ for all flavors. As a consequence, the strangeness quantum number is not conserved in the star. There are therefore only **two conserved charges** with their associated chemical potentials – the baryon number and electric charge.

As a consequence, neutron stars are not made of pure neutrons, as originally proposed. Some neutrons will beta-decay until equilibrium between neutrons, protons

and electrons is reached. This is expressed by the relationship $\mu_p = \mu_n - \mu_e$. Since the chemical potentials grow with increasing density, other thresholds are reached, and additional particle species are populated (Fig. 6.16). In a Fermi gas model, the thresholds are simply the masses of the particles. For example, the Λ hyperon will be in equilibrium in the star when the neutron chemical potential attains the value of the lambda mass ($m_\Lambda = 1113$ MeV). In this respect, the isospin symmetry energy arising from the coupling of the baryons isospin to the neutral rho meson is very important. It favors conversion of neutrons to baryons of opposite isospin. For this reason, one has to work with the complicated Lagrangian of field theory discussed in the previous section.

The particle composition is shown in Fig. 6.16. At densities below saturation, the charge-neutral matter is almost pure neutrons with a small admixture of protons and electrons. With increasing density, the electron Fermi energy increases to the muon mass, and the muons will be populated. Hyperon thresholds are reached at about three times nuclear density. At higher density, these hyperons build an important population in neutron star matter.

The above neutron star models show that the appearance of hyperons is connected with the increasing density in neutron star interiors. Bednarek and Manka [67] have recently discussed an equation of state of neutron star matter including strange baryons. The effects of the strength of hyperon–hyperon interactions on the equations of state have been analyzed. These calculations indicate that the change of the

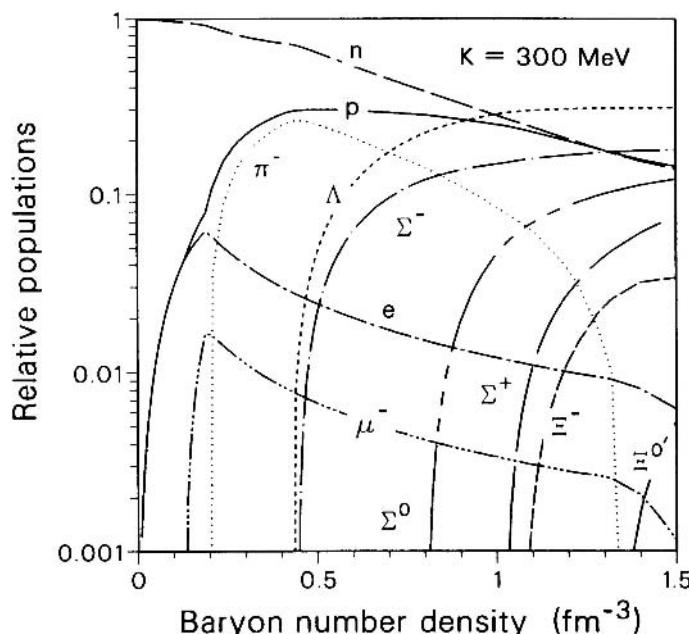


Fig. 6.16. Population of various particles in hadronic stars as a function of baryon density. Nuclear density corresponds to 0.153 fm^{-3} . Figure adapted from Glendenning [5]

hyperon–hyperon coupling constants affects the chemical composition of a neutron star. The corresponding numerical hyperon star models exclude a large population of strange baryons in the stellar interior.

Masses and Radii for Hadronic Stars

As with white dwarfs we can now build sequences of solutions of neutron stars by integrating the TOV equations with the central density as parameter. The central density of stable neutron stars lies between about half and 10 times nuclear energy density ($2.5 \times 10^{14} \text{ g cm}^{-3}$). In Fig. 6.12 we compare the equation of state for pure neutron matter and stable neutron matter including hyperons (H). With the addition of a new baryon species, the EoS is softened, since the Fermi pressure of neutrons and protons near the top of their Fermi seas is relieved by allowing them to hyperonize to unoccupied low-momentum states. For these EoS, the TOV equations can be integrated to yield the masses and radii as function of the central density (Fig. 6.17).

Due to the softening by hyperons, neutron stars containing hyperons are more compact and reach a lower upper mass limit. Remember that stability is lost when the slope in $M(\rho_c)$ becomes negative. Below the minimum mass of neutron stars and for all lower central densities down to the white dwarf maximum mass, the fundamental vibrational mode is unstable. If a stable neutron star below the mass limit were to accrete matter so as to move beyond the mass limit, the unstable fundamental vibrational mode would destroy it, and the object would collapse to form a black hole. This upper mass limit is around $1.5 M_\odot$ for neutron stars based

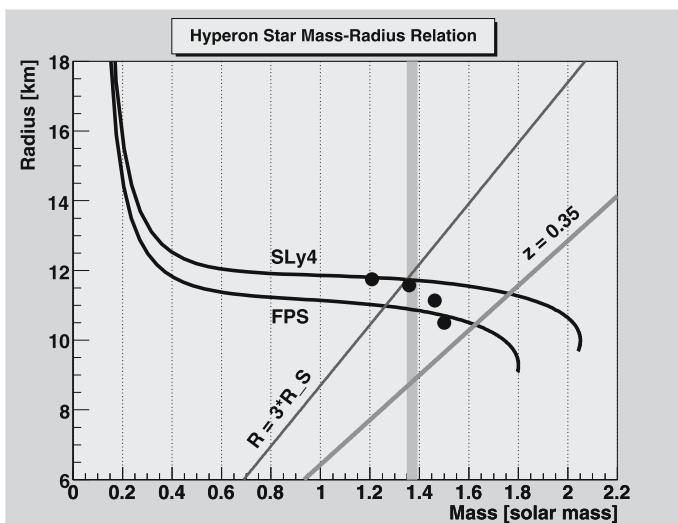


Fig. 6.17. Neutron star radii as a function mass for pure neutron matter vs. hyperon matter (dots). The upper mass limit is lowered by hyperonization

on hyperonization. For that reason it has been argued that the observed masses $M \simeq 1.4 M_{\odot}$ are in fact near the upper mass limit.

In Fig. 6.17, the mass-radius relation is shown for hadronic stars. It is a remarkable fact that the radius stays roughly constant in the intermediate mass range $0.5 \leq M \leq 1.3 M_{\odot}$ with typical values of $R \simeq 13.5$ km.

6.3.2 Quark Matter Cores

If a neutron star gets dense enough, the quarks inside the neutrons and protons are liberated and a cold quark liquid forms. This form of matter has remarkable similarities to the state of electrons in a metal. This phenomenon can lead to superconductive states, for quark matter this phase is called **color-superconductivity**.

Phase Diagrams

In principle, we should be able to calculate the behavior of high-density matter. We have a theory for strong interaction, the QCD, but it is difficult to work with this theory in the low density limit.

Along the horizontal axis of Fig. 6.18, the temperature is zero, and density rises from the onset of nuclear matter through a transition to quark matter. Compact stars are found in this region. Along the vertical axis, the temperature rises. At low temperature, we find the hadronic gas, in which the quarks are confined into

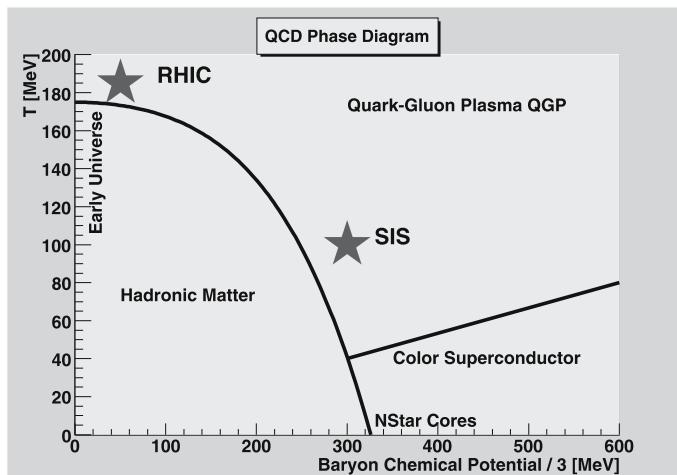


Fig. 6.18. Phase diagram for two light quarks and a strange quark. At sufficiently high densities cold quark matter is in the CFL phase, in which quarks of all three colors and flavors form Cooper pairs. RHIC and SIS are the names of heavy-ion collider experiments, which probe QCD at high temperature. While heavy-ion collider experiments can only probe the phase transition at high temperature, the cores of neutron stars are sensitive to the low temperature region, where the 2SC and CFL phase are probably the most important ones

neutrons and protons. At higher temperatures, this phase makes a transition to the quark-gluon plasma (QGP), in which quarks and gluons are unconfined. This is the region explored by high-energy heavy-ion colliders RHIC at Brookhaven, LHC at CERN and SIS at GSI (Darmstadt).

The low-temperature region is where quark matter will occur in neutron stars. One expects that some of the quarks will pair up, forming Cooper pairs, similar to electrons in metals. Cooperic pairing is a generic phenomenon of fermions: electrons and quarks are both fermions, with some **attractive interaction** between them.

A Few Words on Superconductivity

The understanding of superconductivity was advanced in 1957 by three American physicists John Bardeen, Leon Cooper, and John Schrieffer, through their theories of superconductivity, known as the BCS theory. The BCS theory explains superconductivity at temperatures close to absolute zero. Cooper realized that atomic lattice vibrations were directly responsible for unifying the entire current. They forced the electrons to pair up into teams that could pass all of the obstacles which caused resistance in the conductor. These teams of electrons are known as Cooper pairs. Cooper and his colleagues knew that electrons which normally repel one another must feel an overwhelming attraction in superconductors. The answer to this problem was found to be in phonons, packets of sound waves present in the lattice as it vibrates.

According to the theory, as one negatively charged electron passes by positively charged ions in the lattice of the superconductor, the lattice distorts. This in turn causes phonons to be emitted which forms a trough of positive charges around the electron. Before the electron passes by and before the lattice springs back to its normal position, a second electron is drawn into the trough. It is through this process that two electrons, which should repel one another, link up. The forces exerted by the phonons overcome the electrons' natural repulsion. The electron pairs are coherent with one another as they pass through the conductor in unison. The electrons are screened by the phonons and are separated by some distance. When one of the electrons that make up a Cooper pair passes close to an ion in the crystal lattice, the attraction between the negative electron and the positive ion cause a vibration to pass from ion to ion until the other electron of the pair absorbs the vibration. The net effect is that the electron has emitted a phonon and the other electron has absorbed the phonon. It is this exchange that keeps the Cooper pairs together. It is important to understand, however, that the pairs are constantly breaking and reforming. Because electrons are indistinguishable particles, it is easier to think of them as permanently paired.

Fermions are particles that obey the Pauli exclusion principle, which says that no two fermions can be in the same state. So as you add more and more fermions to a finite-sized box, you have to put them in higher and higher momentum states. For noninteracting fermions at zero temperature you would just end up with a “Fermi sea” of filled states: all states with energy less than the Fermi energy $E_F = \mu$ are filled, and all states above E_F are empty. But if there is an attractive interaction

between the fermions, then things are very different. There is a state of lower free energy than a simple Fermi surface. That state is called the “BCS” state, named after Bardeen, Cooper, and Schrieffer, who first identified it, and it arises from a complicated quantum-coherent superposition of Cooper pairs of particles (and holes).

In the case of electrons, their dominant interaction is electrostatic repulsion, and it is only the presence of a background lattice of positively charged ions that allows additional attractive phonon-mediated interactions to exist. The resultant Cooper pairing is rather fragile, and easily disrupted by thermal fluctuations, hence metals only become superconducting at very low temperatures. The condensate of Cooper pairs of electrons is charged, and as a result the photon, which couples to electric charge, becomes massive. Superconducting metals therefore contain neither electric nor magnetic fields. A perfect conductor cannot contain electric fields (the charges would rearrange themselves to cancel it), but the special thing about a superconductor is that it expels magnetic fields (known as the Meissner effect).

Color Superconductivity of Quarks

For quarks things are very different. The dominant interaction between quarks is the strong interaction, described by QCD, which is very attractive in some channels (after all, QCD binds quarks together to form baryons). This leads us to expect that quarks will form Cooper pairs very readily and that **quark matter will generically acquire a condensate of Cooper pairs**. Since pairs of quarks cannot be color-neutral, the resulting condensate will break the local color symmetry, making the gluons massive. We call this “color superconductivity.” Note that the quark pairs play the same role here as the Higgs particle does in the standard model: the color-superconducting phase can be thought of as the Higgs phase of QCD.

Color Superconducting Phases

Color-superconducting quark matter can come in a rich multiplicity of different possible phases, based on different pairing patterns of the quarks. This is possible, because quarks come in three different colors, and at the density of a compact star core we expect three different flavors: up, down, and strange. Recent work has concentrated on calculating which type of pairing is favored at which density. This is a complicated problem, in which one must take into account the requirement that bulk matter be neutral with respect to both electric and color charge, as well as equilibration under the weak interaction processes that can turn one quark flavor into another, and finally the strange quark mass. The results so far, starting at the highest densities and working down, are roughly this:

- Highest densities: **color-flavor-locked** (CFL) quark pairing, in which all three flavors participate symmetrically. CFL quark matter has many special properties, including the fact that chiral symmetry is broken by a new mechanism: the quark pairs themselves, instead of the more conventional chiral condensate $\langle \bar{q}q \rangle$. There may be kaon condensation.

- Very high densities: **Gapless CFL**, in which holes start to open up in the CFL pairing pattern, leaving some quarks unpaired.
- Middle high densities: unknown. Many possibilities have been suggested, including crystalline pairing, two-flavor pairing (2SC), single-flavor pairing, color-spin locking, etc.

We know what phase is favored in the limit of infinite density (CFL phase), but the nature of the pairing in quark matter at realistic neutron-star densities is still a vigorously debated question.

Unpaired Quarks at High Densities

At densities of about twice nuclear density hyperons appear in neutron star matter. This is called the **hadronic phase** (HP) of neutron star matter. At densities much above the nuclear density this hadronic phase is allowed to undergo a phase transition to the **deconfined quark matter phase** (QP). Quark matter consists of u , d and s quarks, as well as electrons in weak equilibrium

$$d \rightarrow u + e^- + \bar{\nu}_e \quad (6.110)$$

$$s \rightarrow u + e^- + \bar{\nu}_e \quad (6.111)$$

$$s + u \leftrightarrow d + u . \quad (6.112)$$

Quark matter is therefore described in terms of four chemical potentials μ_u , μ_d , μ_s and μ_e which must satisfy the relations

$$\mu_s = \mu_d = \mu_u + \mu_e . \quad (6.113)$$

The chemical potential of neutrinos vanishes, since neutrinos can escape from the interior of the star. This relation reduces therefore the number of chemical potentials to two independent ones. It is appropriate to choose the pair (μ_n, μ_e) with

$$\mu_n \equiv \mu_u + 2\mu_d = 3\mu_u + 2\mu_e , \quad (6.114)$$

or

$$\mu_u = \frac{1}{3} (\mu_n - 2\mu_e) \quad (6.115)$$

and

$$\mu_s = \mu_d = \frac{1}{3} (\mu_n + \mu_e) . \quad (6.116)$$

In a **pure quark phase** one has to require the charge-neutrality of the quark phase

$$\varrho_c^{QP} = \frac{2}{3}n_u - \frac{1}{3}n_d - \frac{1}{3}n_s - n_e = 0 . \quad (6.117)$$

This fixes the chemical potential of the electrons in terms of the chemical potential μ_n . The EoS is then parametrized by only one chemical potential, say μ_n .

To calculate the EoS of the QP the *effective mass bag model* is applied [355], which is based on the MIT bag model for hadrons [119]. The Lagrangian of quantum chromodynamics (QCD) can be solved only numerically on a lattice and only under conditions which are not relevant to the cold matter of neutron stars. One can gain however some insights using a model of quark confinement in terms of the MIT bag model which was invented to understand the hadronic masses in terms of their quark constituents. The true vacuum of QCD is assumed to have the property that free quarks are excluded. Within the hadronic volume, however, the vacuum must be expelled, and this costs energy. This energy is called the **bag constant** and is denoted by B . The presence of quarks in a volume V leads therefore to an energy BV . In addition there is also kinetic energy associated with the presence of quarks in a volume. For hadrons it is assumed that the quarks can move freely in a spherical volume. For quark matter, energy associated with boundaries is unimportant compared to the energy content of the interior: we can assume that the quarks are just fermions. Therefore, the energy density and pressure of quark matter will consist of the kinetic motion of the quarks and the contribution of the confining bag. The optimum value of the bag constant B that can account for hadronic masses is $B^{1/4} = 145$ MeV. This might however change for quark matter so that we consider B as a free constant of the model. In fact, B must exceed a certain minimum value in order that the energy per baryon number fits the value of iron.

In the MIT bag model, the quarks assume an effective mass by means of their interaction with a gluon background (one-loop approximation)

$$m_q^*(\mu_q) = \frac{m_q}{2} + \sqrt{\frac{m_q^2}{4} + \frac{g^2 \mu_q^2}{6\pi^2}}, \quad (6.118)$$

with the quark mass m_q of flavor q and the strong coupling constant g . Similarly to what we have found in the mean field theory of nucleon matter, baryon number density n , energy density ϵ and pressure P are given by the expressions

$$n = \frac{1}{3} \sum_q \frac{\mathcal{N}_q}{2\pi^2} \int_0^\infty k^2 dk [n(k, \mu_q) - n(k, -\mu_q)] \quad (6.119)$$

$$\epsilon = \sum_q \frac{\mathcal{N}_q}{2\pi^2} \int_0^\infty E_q k^2 dk [n(k, \mu_q) + n(k, -\mu_q)] + B \quad (6.120)$$

$$P = \frac{1}{3} \sum_q \frac{\mathcal{N}_q}{2\pi^2} \int_0^\infty k \frac{\partial E_q(k)}{\partial k} k^2 dk [n(k, \mu_q) + n(k, -\mu_q)] - B \quad (6.121)$$

where

$$E_q(k) = \sqrt{m_q^2 + k^2} \quad (6.122)$$

is the quark kinetic energy, and

$$n(k, \mu_q) = \frac{1}{\exp[(E_q(k) - \mu_q)/k_B T] + 1} \quad (6.123)$$

is the Fermi distribution for temperature T . The sum extends over all quark flavors with degree of freedom $\mathcal{N}_q = 2 \times 3$ for two spin states and three color states. The factor $1/3$ occurs in the density, since there are three quarks per baryon. The bag constant shifts the energy by a positive amount per unit volume in the deconfined state relative to the true vacuum, but it shifts the pressure by a negative amount, since a vacuum energy represents a negative pressure.

These expressions can be evaluated in the **zero-temperature limit**, since the distribution functions become step functions with k_q as the Fermi momentum of flavor q . One has just to remember the integral for the calculation of the pressure

$$\begin{aligned} & \int_0^{k_q} \frac{k^4 dk}{\sqrt{m^2 + k^2}} \\ &= \frac{1}{4} \left[k_q^3 \sqrt{k_q^2 + m^2} - \frac{3}{2} m^2 k_q^2 \sqrt{m^2 + k_q^2} + \frac{3}{2} m^4 \ln \left(\frac{\sqrt{k_q^2 + m^2} + k_q}{m} \right) \right]. \end{aligned} \quad (6.124)$$

With this one finds

$$n = \sum_q \frac{k_q^3}{3\pi^2} = \frac{1}{3}(n_u + n_d + n_s) \quad (6.125)$$

$$\epsilon = B + \sum_q \frac{3}{4\pi^2} \left[\mu_q k_q \left(\mu_q^2 - \frac{1}{2} m_q^2 \right) - \frac{1}{2} m_q^4 \ln \left(\frac{\mu_q + k_q}{m_q} \right) \right] + \epsilon_e \quad (6.126)$$

$$P = -B + \frac{1}{3} \sum_q \frac{3}{4\pi^2} \left[\mu_q k_q \left(\mu_q^2 - \frac{5}{2} m_q^2 \right) + \frac{3}{2} m_q^4 \ln \left(\frac{\mu_q + k_q}{m_q} \right) \right] + P_e. \quad (6.127)$$

The Fermi momentum is defined in terms of the chemical potential

$$\mu_q = \sqrt{m_q^2 + k_q^2} \quad (6.128)$$

with masses $m_u = 5$ MeV, $m_d = 7$ MeV and $m_s = 150$ MeV. The other three quarks have masses exceeding 1.5 GeV, so they are not contributing for neutron star matter. For u and d quarks we can simply write

$n_q = \frac{\mu_q^3}{\pi^2 (\hbar c)^3}$

(6.129)

and for strange quarks

$$n_s = \frac{\mu_s^2 - m_s^2 c^4}{\pi^2 (\hbar c)^3} \sqrt{\mu_s^2 - m_s^2 c^4}. \quad (6.130)$$

Since the electrons are highly relativistic, their contribution is simply

$$\epsilon_e = \frac{\mu_e^4}{4\pi^2 (\hbar c)^3}, \quad P_e = \frac{\mu_e^4}{12\pi^2 (\hbar c)^3}. \quad (6.131)$$

For vanishing quark masses, all three flavors have the same density, $n_u = n_d = n_s$ and $n_e = 0$. With increasing mass of the strange quark, the number density of strange quarks decreases with respect to the other two flavors, but the electron density increases in order to compensate the charge neutrality. In leading order, the pressure can simply be approximated by

$$P_Q \simeq \frac{N_f N_c}{12\pi^2} \mu^4 - B_{\text{eff}}. \quad (6.132)$$

The energy density follows from this by

$$\epsilon_Q = \mu \frac{\partial P_Q}{\partial \mu} - P_Q = \frac{N_f N_c}{4\pi^2} \mu^4 + B_{\text{eff}}. \quad (6.133)$$

For massless quarks we obtain therefore a very simple equation of state

$$P_Q = \frac{1}{3}(\epsilon_Q - 4B). \quad (6.134)$$

This relation is only slightly changed for nonvanishing strange quark mass. It can also be improved by adding lowest order corrections due to interaction, in leading order one obtains

$$P_{Q,1st} = -\frac{\alpha_s N_f (N_c^2 - 1)}{16\pi^3} \mu^4 + O(\alpha_s^2), \quad (6.135)$$

where $\alpha_s = g^2/4\pi$ is the value of the running coupling constant of strong interaction defined at the scale of the quark chemical potential. This correction can be included into the pressure equation by defining a parameter c [31]

$$P_Q \simeq \frac{N_f N_c}{12\pi^2} \mu^4 (1 - c) - B_{\text{eff}}. \quad (6.136)$$

The chemical potential of the electrons is very small in pure quark matter, $\mu_e \simeq 22$ MeV. In these expressions, B is considered to be a free parameter ranging from $B^{1/4} = 165$ MeV ($B = 96$ MeV fm $^{-3}$) to $B^{1/4} \simeq 200$ MeV ($B \simeq 208$ MeV fm $^{-3}$). The lower bound is given if one requires that the deconfinement phase transition does not occur at densities below nuclear density.

In the following we give a short outline of the properties of Cooper paired quark matter (for more details, see, e.g. the lecture notes by Shovkovy [368]).

6.3.3 Grand Canonical Potential for Quark Matter

Unpaired Quarks

For systems with changing particle numbers the grand canonical potential $\Omega(V, T, \mu)$ is the most suitable quantity, thermodynamically given by

$$\Omega(V, T, \mu) = E(V, T, \mu) - S(V, T, \mu)T - N(V, T, \mu)\mu. \quad (6.137)$$

In statistical mechanics, Ω for Bose or Fermi systems is given by the grand partition sum

$$\Omega(V, T; \mu) = \pm k_B T \sum_i \ln \left[1 \mp \exp \left(\frac{\mu - \sqrt{m^2 + p^2}}{k_B T} \right) \right]. \quad (6.138)$$

For particles filling a phase-space volume $d^3q d^3p$, the sum has to be replaced in terms of the volume integration

$$\sum_i \rightarrow \frac{g}{(2\pi\hbar)^3} \int d^3q d^3p \quad (6.139)$$

with the replacement of the energy levels $\epsilon_i \rightarrow \epsilon = \sqrt{m^2 + p^2}$ and the number of degrees of freedom g . For homogeneous Fermi systems, we obtain therefore the following expression

$$\Omega(V, T, \mu) = -\frac{gV}{2\pi^2\hbar^3} k_B T \int_0^\infty dp p^2 \ln \left[1 \mp \exp \left(\frac{\mu - \sqrt{m^2 + p^2}}{k_B T} \right) \right]. \quad (6.140)$$

From this we get the expression for the pressure

$$P(T, \mu) = -\frac{\partial \Omega}{\partial V}, \quad (6.141)$$

or for homogeneous Fermi systems

$$\Omega(V, T, \mu) = -P(T, \mu)V. \quad (6.142)$$

The internal energy density ϵ follows from

$$\epsilon(T, \mu) = \frac{E}{V} = \frac{\Omega + S T + N \mu}{V} = -P + \frac{\partial P}{\partial T}|_\mu T + \frac{\partial P}{\partial \mu}|_T \mu, \quad (6.143)$$

i.e.

$$\boxed{\epsilon(T, \mu) = -P(T, \mu) + s(T, \mu)T + n(T, \mu)\mu.} \quad (6.144)$$

Here, the particle density is given by

$$\boxed{n(T, \mu) = -\frac{1}{V} \frac{\partial \Omega}{\partial \mu}|_{T,V}.} \quad (6.145)$$

For homogeneous system, it is therefore useful to replace $\Omega \rightarrow \Omega/V$.

Low-Temperature Limit

For a system consisting of three quark flavors we may write

$$\Omega(T, \mu) = - \sum_f \frac{g_f k_B T}{2\pi^2 \hbar^3} \int_0^\infty dp p^2 \ln \left[1 + \exp \left(\frac{\mu_f - \sqrt{m_f^2 + p^2}}{k_B T} \right) \right] + B_{\text{eff}} . \quad (6.146)$$

For $T \rightarrow 0$, we can expand the logarithm and find, $g_f = 2 \times 3$ for quarks,

$$\Omega(T, \mu) = - \sum_f \frac{3}{\pi^2 (\hbar c)^3} \int_0^{\epsilon_F} dp p^2 \left[\mu_f - \sqrt{m_f^2 + p^2} \right] + B_{\text{eff}} . \quad (6.147)$$

For unpaired quarks, we can write therefore in the cold limit

$$\Omega(\mu_u, \mu_d, \mu_s) = - \frac{3}{\pi^2 (\hbar c)^3} \sum_{f=u,d,s} \int_0^{\sqrt{\mu_f^2 - m_f^2}} dp p^2 \left[\mu_f - \sqrt{m_f^2 + p^2} \right] + B_{\text{eff}} , \quad (6.148)$$

where $\mu_u = \mu - 2\mu_e/3$ and $\mu_d = \mu_s = \mu + \mu_e/3$ are the individual quark chemical potentials, μ are the baryon and μ_e the electron chemical potential, respectively. The masses of the up and down quarks can be neglected, since $m_u \simeq m_d \ll \mu$. The strange quark mass is however not negligible with respect to μ . For neutron stars, it is possible to have $m_s < \mu$, so that we can expand in m_s/μ .

In neutral unpaired quark matter, the electron chemical potential is determined by the condition of charge neutrality, i.e.

$$\mu_e = \frac{m_s^2}{4\mu} - \frac{m_s^4}{48\mu^3} + O(m_s^6/\mu^5) . \quad (6.149)$$

Substituting this into the expression for the potential we obtain

$$\begin{aligned} \Omega_{\text{unp}}(\mu) = & - \frac{3}{4\pi^2} (1 - c) \mu^4 + \frac{3\mu^2 m_s^2}{4\pi^2} + (12 \ln(m_s/2\mu) - 7) \frac{m_s^4}{32\pi^2} \\ & + \frac{5m_s^6}{576\pi^2 \mu^2} + B_{\text{eff}} + O(m_s^8/\mu^4) . \end{aligned} \quad (6.150)$$

Here we have introduced the parameter c which corrects for higher orders in the quark running coupling constant. This expansion is in fact rapidly convergent in m_s/μ , even when $m_s = 300$ MeV and $\mu = 350$ MeV.

Color-Flavor-Locked (CFL) Quark Matter

In the CFL phase, the pairing locks the Fermi momenta of all the quarks to a single value. This costs free energy, which is offset by the pairing contribution

$$\Omega_\Delta = -\frac{3}{\pi^2} \Delta^2 \mu^2 + O(\Delta^4). \quad (6.151)$$

The calculation of Δ yields values in the range of 10–100 MeV for μ in the range of 300–600 MeV (Rischke [343]). When color and electric neutrality is imposed, we do not find any electrons in the CFL phase, since there are equal numbers of up, down and strange quarks, i.e. $\mu_e = 0$. In this sense, we obtain

$$\Omega_{\text{CFL}} = \Omega_{\text{unp}} + \frac{3m_s^4 - 48\Delta^2\mu^2}{16\pi^2}. \quad (6.152)$$

This shows that all such models can be accommodated into an expansion of the form

$$\Omega_{\text{CFL}}(\mu) = -\frac{3}{4\pi^2}(1-c)\mu^4 + \frac{3a_2}{4\pi^2}\mu^2 + B_{\text{eff}}, \quad (6.153)$$

with the coefficient

$$a_2 = m_s^2 - 4\Delta^2. \quad (6.154)$$

In leading order, pressure is given by

$$P = \frac{3\mu^4}{4\pi^2}(1-c) - B_{\text{eff}}, \quad (6.155)$$

the energy density by

$$\epsilon = 3(1-c)\frac{3\mu^4}{4\pi^2} + B_{\text{eff}} \quad (6.156)$$

and therefore the EoS by

$$P = \frac{1}{3}(\epsilon - 4B_{\text{eff}}) \quad (6.157)$$

completely independent of the correction given by c . The EoS of quark matter has a threshold in energy density given by $4B_{\text{eff}}$ (see Fig. 6.20).

Gapless CFL Phases

The blue up and down quarks do not participate in Cooper pairing in the 2SC phase. They give rise to gapless quasiparticles in the low energy spectrum of the theory. The presence of the ungapped blue quarks could result in a large neutrino emissivity due to beta-processes, $d_b \rightarrow u_b + e^- + \bar{\nu}_e$ and $u_b + e^- \rightarrow d_b + \nu_e$. The other four quarks (red and green) are gapped.

The pressure of the cold dense quark matter is dominated by the Pauli exclusion pressure. The partial contribution of each quark is $\mu^4/12\pi^2$ (see above). The contribution due to diquark pairing is given by $(\mu\Delta/2\pi)^2$ per each gapped quark

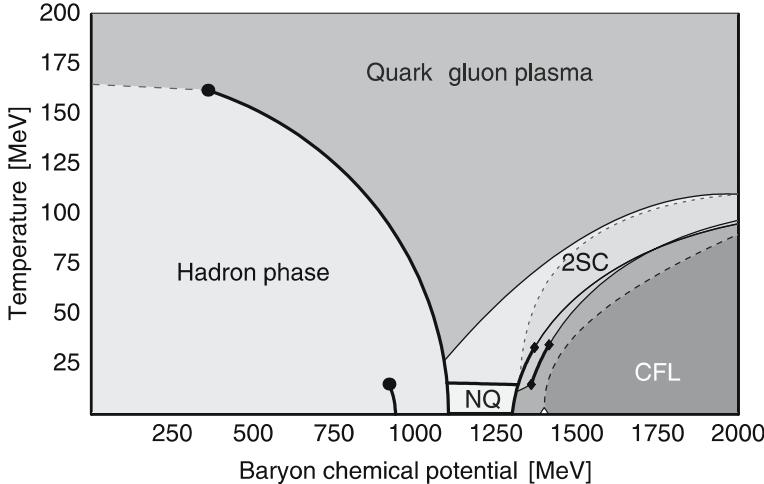


Fig. 6.19. Phase diagram for neutral quark matter within the field-theoretical Nambu–Jona–Lasinio model [348], see also [349, 350]. The *thin solid lines* mark second-order phase boundaries between two phases that differ by one or more diquark condensates. The *dashed lines* mark the (dis-)appearance of gapless modes in different phases. First-order phase boundaries are denoted by *heavy lines*. Figure adapted from [348]

quasiparticle. In the 2SC phase, there are six (two flavors and three colors) quarks in total, and four of them give rise to quasiparticles with gap Δ in their energy spectra. Thus the pressure is approximately given by

$$P_{2SC} \simeq \frac{N_c N_f}{12\pi^2} \mu^4 + 4 \left(\frac{\mu \Delta}{2\pi} \right)^2 - B_{\text{eff}} = \frac{\mu^4}{2\pi^2} + \frac{\mu^2 \Delta^2}{\pi^2} - B_{\text{eff}}. \quad (6.158)$$

From this, we can derive the corresponding energy density

$$\epsilon_{2SC} \simeq \frac{3\mu^4}{2\pi^2} + \frac{\mu^2 \Delta^2}{\pi^2} \left(1 + \frac{2\mu}{\Delta} \frac{\partial \Delta}{\partial \mu} \right) + B_{\text{eff}}. \quad (6.159)$$

These two expressions give the parametric representation of the EoS of dense quark matter in the 2SC phase in lowest order of approximation.

CFL Paired Color-Superconducting Quark Phases

At high densities, strange quarks will also contribute. The constituent strange quark mass in vacuum QCD is estimated to be of order 500 MeV. Its current mass is however only $\simeq 100$ MeV. For dense matter with say $\mu \simeq 500$ MeV, the value of the strange quark mass should be in the range of 100 to 500 MeV. It is then plausible that strange quarks also participate in Cooper pairing. This can be discussed in the

limit, where all three quarks are massless. In this case, the quark model possesses a global $SU(3)_L \times SU(3)_R$ chiral symmetry and the global $U(1)_B$ symmetry related to baryon number conservation. This is in addition to the $SU(3)_c$ color gauge symmetry of QCD.

The favored pairing pattern at high densities, where the strange quark Fermi momentum is close to the up and down Fermi momenta, is color-flavor-locking (CFL). This has now been confirmed by various calculations (NJL models and gluon-mediated models, Fig. 6.19). The CFL spin-0 condensate corresponds to the following ground-state expectation value

$$\langle q_i^\alpha q_j^\beta \rangle \propto C\gamma_5 \left((\kappa + 1)\delta_i^\alpha \delta_j^\beta + (\kappa - 1)\delta_j^\alpha \delta_i^\beta \right). \quad (6.160)$$

Color indices α, β and flavor indices i, j run from 1,2,3. Dirac indices are suppressed and C is the Dirac conjugation matrix. The parameter κ is small, but not zero. The Kronecker deltas connect color indices with flavor ones, so that the condensate is not invariant against color rotations, nor under flavor rotations, but only under simultaneous color and flavor rotations. The feature of the CFL phase are then:

- The color gauge group $SU(3)_C$ is broken. All eight gluons become massive.
- All the quark modes are gapped. The nine quasiquarks fall into $8 \otimes 1$ of the unbroken global $SU(3)$. There are therefore two gap parameters, one for the singlet state and one for the octet state.
- Electromagnetism survives unbroken (photon + one gluon).
- Two global symmetries are broken: the chiral symmetry and baryon number.

In a real star, we must require electromagnetic and color neutrality.

In contrast to the 2SC phase, there are no gapless quark quasiparticles. Therefore, at small temperatures, $k_B T < \Delta$, the quasiparticles determine all transport properties in quark matter. Unlike the 2SC phase, the CFL phase is superfluid, it is however not an electromagnetic superconductor (electromagnetism is unbroken). The Pauli pressure is built up by nine quarks (three flavors and three colors) with one quasiparticle gap Δ_1 and eight quasiparticles with gap Δ_2

$$\begin{aligned} P_{\text{CFL}} &\simeq \frac{N_c N_f}{12\pi^2} \mu^4 + \left(\frac{\mu \Delta_1}{2\pi} \right)^2 + 8 \left(\frac{\mu \Delta_2}{2\pi} \right)^2 - B_{\text{eff}} \\ &\simeq \frac{3\mu^4}{4\pi^2} + 3 \frac{\mu^2 \Delta^2}{\pi^2} - B_{\text{eff}}, \end{aligned} \quad (6.161)$$

where we used the approximate relation between the singlet and octet gaps, $\Delta_1 = 2\Delta_2 = 2\Delta$. From this, one can derive the corresponding energy density

$$\epsilon_{\text{CFL}} \simeq \frac{9\mu^4}{4\pi^2} + 3 \frac{\mu^2 \Delta^2}{\pi^2} \left(1 + \frac{2\mu}{\Delta} \frac{\partial \Delta}{\partial \mu} \right) + B_{\text{eff}}. \quad (6.162)$$

As we have shown above, a different derivation for the EoS of quark matter can be based on the grand canonical potential \mathcal{Q} for fermionic systems. We have

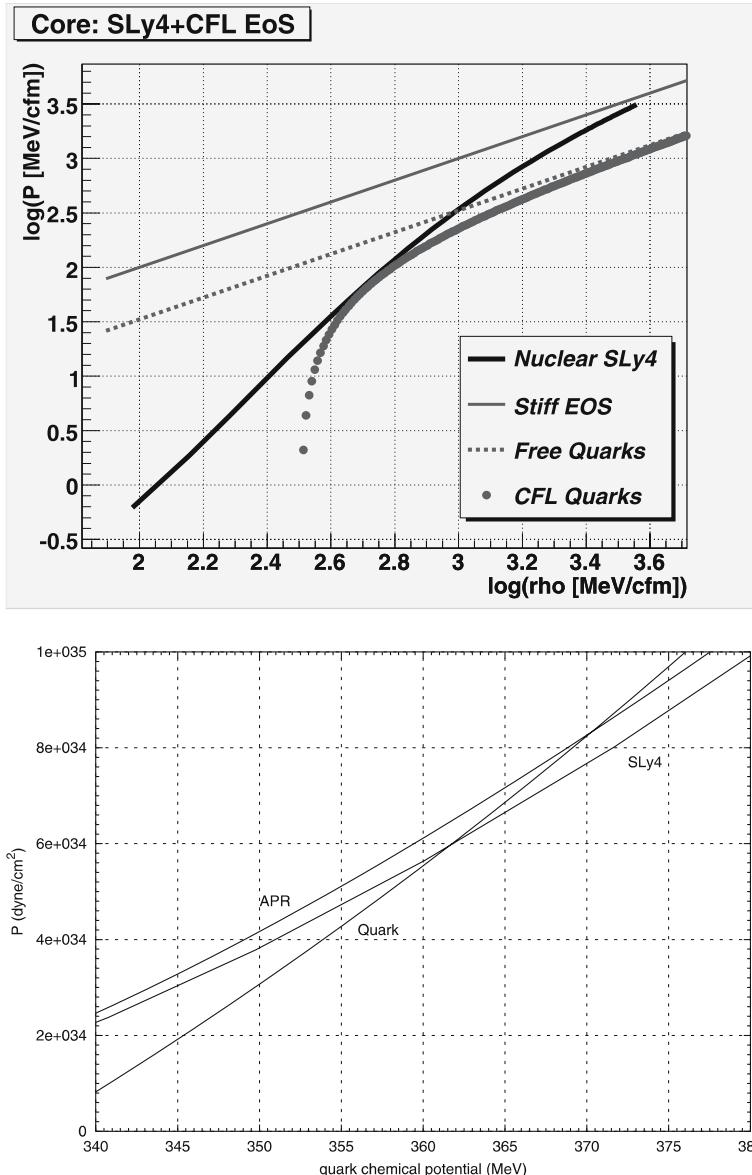


Fig. 6.20. EoS for outer and inner core with the quark–hadron phase transition for a stiff nuclear EoS (SLy4). *Top:* A continuous transition is obtained for $m_s = 2\Delta$ and a low bag constant of $B = 80 \text{ MeV fm}^{-3}$ with a transition density of $\simeq 3.3$ nuclear density. The *upper solid line* corresponds to the EoS $P = \varrho$, the *lower dotted line* to $P = \varrho/3$ for asymptotically free quarks. *Bottom:* Maxwell construction for the phase transition with the two nuclear EoS APR and SLy4. SLy4 enables quark cores at lower density than APR

shown that quark matter in the CFL phase is always given by (see also Alford et al. [31, 32])

$$\Omega_{\text{CFL}}(\mu) = -\frac{3}{4\pi^2}(1-c)\mu^4 + \frac{3a_2}{4\pi^2}\mu^2 + B_{\text{eff}}, \quad (6.163)$$

with

$$a_2 = m_s^2 - 4\Delta^2. \quad (6.164)$$

All pure nuclear EoS (not including hyperons) are by far too stiff beyond three times nuclear density (Fig. 6.20). The above pressure (6.161) nicely extends the nuclear EoS SLy4 beyond about 3.3 times nuclear density for $m_s = 2\Delta$ ($a_2 = 0$) and a relatively low bag constant $B = 80 \text{ MeV fm}^{-3}$. Instead of approaching the stiff EoS $P = \varrho$ of strongly repulsing nuclear matter, quark matter always satisfies the limit $P = \varrho/3$ at high densities. This soft EoS for quark matter will considerably lower the maximum mass of neutron stars.

Neutron Star Models Including Quark Cores

In neutron stars, one expects a transition from the hadronic phase to a 2SC (or CFL) phase not at densities below twice nuclear saturation density. Neutron star models including a possible quark matter phase can be calculated by solving the TOV equation and using BPS EoS at low densities with some modification given by APR98, or SLy4, and a phase transition to a quark matter phase occurring around twice nuclear density. According to the above discussion, we may parametrize the pressure of the quark phase in the following form [31, 32]

$$P_Q = \frac{3}{4\pi^2}\mu^4(1-c) - \frac{3}{4\pi^2}m_s^2\mu^2 + \frac{3}{\pi^2}\Delta^2\mu^2 - B. \quad (6.165)$$

For quark matter consisting of three flavors of noninteracting quarks, $c = 0$. However, once QCD corrections are taken into account, we expect $c \neq 0$. Since QCD corrections are not negligible, they can be subsumed into the parameter c with its most probable value $c \simeq 0.3$. Corrections of order μ^2 in the pressure are due to the nonvanishing strange quark mass m_s and the gap parameter. Inspection of Fig. 6.20 shows that the $c = 0.3$ quark matter EoS is very similar to the nuclear EoS over the pressure range 10 to 200 MeV/fm³.

The corresponding mass-radius relations are shown in Fig. 6.21 for two values of the gap parameter Δ and fixed bag constant and strange quark mass. The solution is given for a bag constant $B^{1/4} = 180 \text{ MeV}$ ($B = 137 \text{ MeV fm}^{-3}$) and a strange quark mass $m_s = 200 \text{ MeV}$ [31, 32]. The core of these neutron star models is made of quark matter and at some radius there is a transition to nuclear matter, which is the favored phase at low densities. The transition pressure is sensitive to Δ . Neutron stars including quark cores are always more compact than pure hadronic models. Secondly, the maximum mass will not exceed 1.7 solar masses, depending on the value of the correction parameter c (Fig. 6.22).

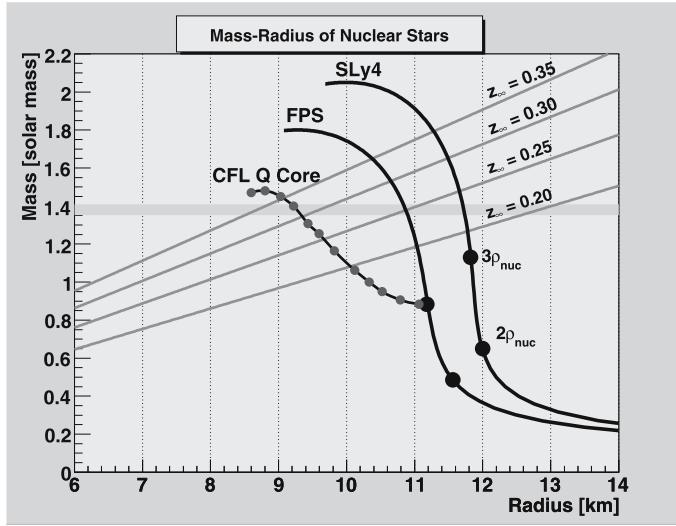


Fig. 6.21. Schematic mass–radius relationship for neutron star models including a quark core compared with pure nuclear matter models. In heavier neutron stars, a quark transition is expected to occur for central densities higher than about three times nuclear densities. The observed redshift of 0.35 [126] would require by far too high masses for normal hadronic stars. A quark core would bring down the mass into the range of observed neutron star masses

APR + bag model QM EoS

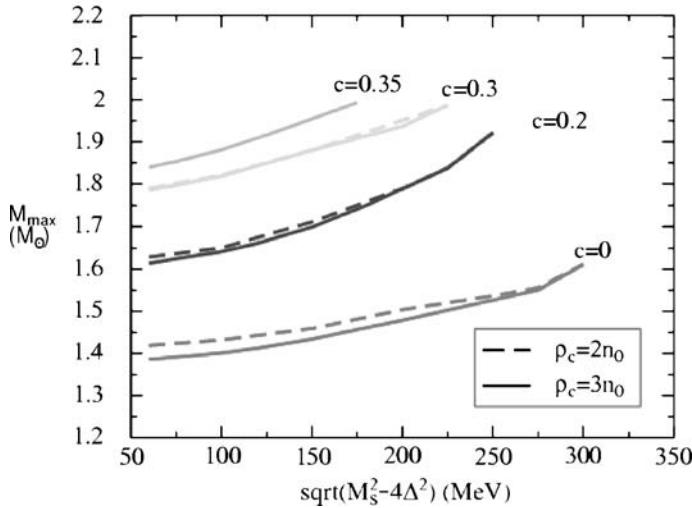


Fig. 6.22. Maximum mass for quark matter stars, as a function of the two parameters m_s and Δ , and c . Figure adapted from Alford et al. [31]

Gravitational Redshift of Neutron Stars

The large collecting area of XMM–Newton has provided, for the first time, detailed spectral information for many of these weak X-ray sources (Fig. 6.23). The EPIC camera can make very fast timing studies of X-ray pulsars, and with its excellent sensitivity, is able to search over much greater distances for such object. Its spatial resolution has also been necessary to separate X-ray pulsars from their surrounding supernova remnants (SNRs).

The fundamental properties of neutron stars provide a direct test of the equation of state of cold nuclear matter, a relationship between pressure and density that is determined by the physics of the strong interactions between the particles that constitute the star. The most straightforward method of determining these properties is by measuring the gravitational redshift of spectral lines produced in the neutron star photosphere. The equation of state implies a mass–radius relation, while a measurement of the gravitational redshift at the surface of a neutron star provides a direct constraint on the mass-to-radius ratio. Cottam et al. [126] have reported the discovery of significant absorption lines in the spectra of 28 bursts of the low-mass X-ray binary EXO 0748–676. Cottam et al. [126] identify the most significant features with the Fe XXVI and XXV $n = 2\text{--}3$ and O VIII $n = 1\text{--}2$ transitions, all with a redshift of $z = 0.35$, identical within small uncertainties for the respective transitions (Fig. 6.23). For an astrophysically plausible range of masses (M approximately 1.3–1.5 solar masses), this redshift is not consistent with standard models of neutron stars composed of normal nuclear matter (hadronic neutron stars), while it could hint to a possibly more compact inner core made of quark matter, as shown in Fig. 6.21. The observed redshift would indicate that this neutron star has a fairly high mass, between 1.6 and 1.8 solar masses, or it is quite compact for a standard mass of 1.4 solar masses. This could be the first evidence for the presence of color-superconducting quark matter inside a massive neutron star.

Previous attempts to measure a neutron star’s redshift focused on a star with an enormous magnetic field. Strong fields, however, induce their own redshift. Since the fields from neutron stars are not precisely known, the magnetic component of the stars’ redshift cannot be clearly separated from the gravitational component. In contrast, the object studied by Cottam’s team has such a weak magnetic field that its redshift results entirely from gravitational effects.

Mixed Phases

Quark matter need however not be homogeneous. In a certain region, hadronic matter might exist together with the quark phase, both phases may be in equilibrium with each other. This condition is familiar in the case of water when solid and liquid phases are in equilibrium with each other. Such a phase is called **mixed phase**. This phase transition occurs at constant pressure. So quark matter might exist in neutron stars in two forms – the core of the star may be composed of pure quark

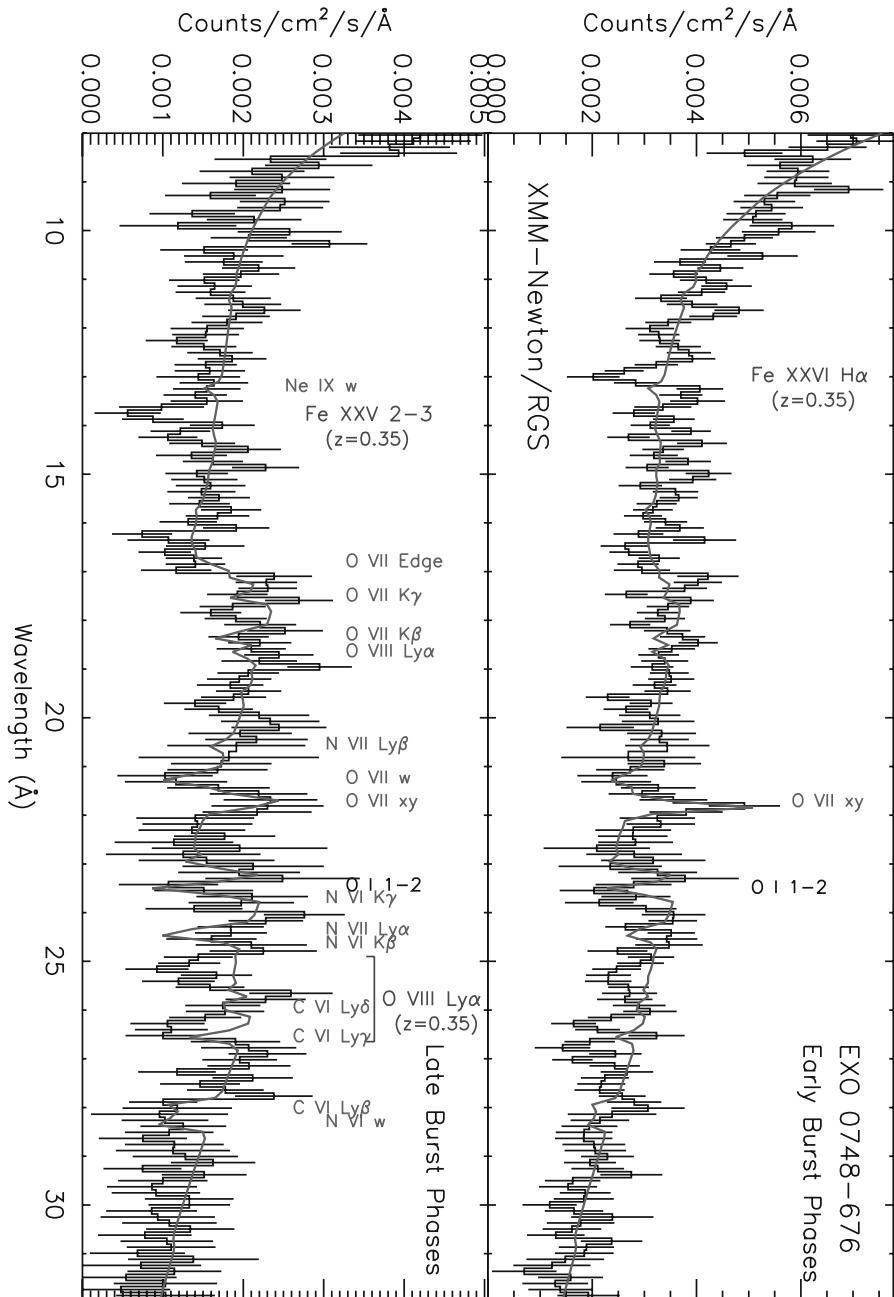


Fig. 6.23. The observed XMM–Newton X-ray spectrum of EXO 0748–676. EXO 0748–676 is a low-mass binary system (LMXB), which shows the famous type X-ray bursts (thermonuclear flashes on neutron star surfaces). Figure adapted from [126]

matter surrounded by a shell of mixed phase in which quark and hadronic matter coexist. Only the lightest stars will contain no quark core and be real hadronic stars.

The essential point for the existence of such a mixed state is that the total charge neutrality can be achieved by a positively charged amount of hadronic matter and negatively charged amount of quark matter. Then we have to deal with two independent chemical potentials, (μ_n, μ_e) . Such a system is usually called a two-component system. The Gibbs condition for mechanical and chemical equilibrium at zero temperature between both phases is just given by the pressure condition

$$P_{HP}(\mu_n, \mu_e) = P_{QP}(\mu_n, \mu_e). \quad (6.166)$$

With this one could calculate the chemical potentials of the mixed phase (MP), where $P_{HP} = P_{QP}$ [355].

6.3.4 Strange Quark Stars

Usually it is assumed that the ground state of hadronic matter is the state in which quarks are confined in individual hadrons. Witten has however proposed that the true ground state of strong interaction is the deconfined state of quark matter consisting of equal numbers of up, down and strange quarks (strange matter hypothesis). From a first point of view this might be incorrect, since what we observe is hadronic matter. But one cannot exclude that the observable Universe is also populated by strange matter objects (sometimes called strangelets). However, soon after this proposal it has been shown that, even if strange matter is the absolute ground state, almost all strange matter would have evaporated in the high-temperature era of the Universe into ordinary hadrons.

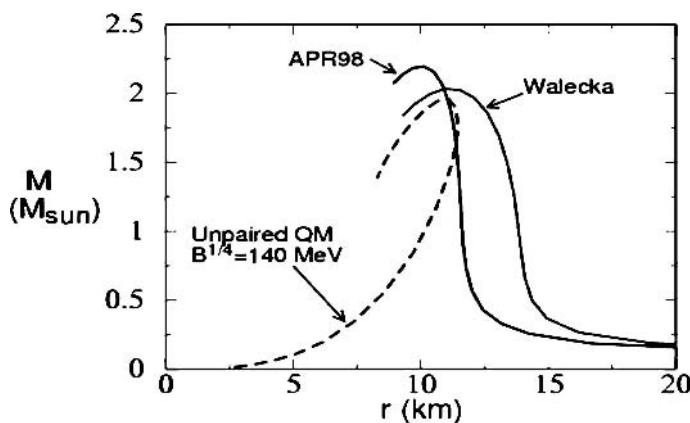


Fig. 6.24. Mass–radius relation for strange quark stars compared to neutron stars

Since the EoS for strange matter is essentially just

$$P = \frac{1}{3}(\varrho c^2 - 4B), \quad (6.167)$$

the pressure vanishes for $\varrho c^2 \rightarrow 4B$. For $B = (145 \text{ MeV})^4$, the density $\varrho_0 = 4 \times 10^{14} \text{ g cm}^{-3}$ is just above nuclear density. In a bound state, the Fermi pressure of the quarks is balanced by the bag pressure. Therefore, for low masses, the mass scales as $M \propto R^3$ (strangelets). For these objects we can calculate the central gravitational pressure

$$P_c = \frac{2\pi G}{3} \varrho_0 R^2. \quad (6.168)$$

When this gravitational pressure exceeds the bag pressure, $4B/3$, gravity will be important. This roughly occurs for $R \simeq 5 \text{ km}$ and masses $M \simeq 0.1 M_\odot$. For higher masses, one has to use the TOV equation together with the above EoS (see Fig. 6.24). Strange stars can also achieve a maximum mass of about (1.5–2.0) M_\odot , depending on the bag constant, but with radii smaller than the radii of standard neutron stars.

6.3.5 The Structure of Massive Neutron Stars

Neutron stars with masses below 1.2 solar masses have central densities not exceeding three times nuclear density. In these stars, the core merely consists of neutrons, protons, electrons and muons. In neutron stars with masses higher than this critical mass, the core has certainly a more complex structure and probably exists of an outer core and an inner one.

The structure of these stars can be summarized as follows (Fig. 6.25):

- In the **outer crust**, matter consists of ions (atomic nuclei) and electrons. The electrons are strongly degenerate, almost ideal Fermi gas, which is relativistic at densities above 10^6 g cm^{-3} . The ions form a strongly coupled Coulomb system, which is solid in most of the crust, but liquid at the lowest densities. The electron Fermi energy grows with increasing density and, as a consequence, nuclei tend to become richer in neutrons. The thickness of the outer crust is a few hundred meters.
- At the base of the outer crust, neutrons begin to drip out of nuclei, thereby producing a neutron gas between the nuclei. This occurs at a density of $4 \times 10^{11} \text{ g cm}^{-3}$. In this **inner crust**, matter consists of electrons, free neutrons and neutron-rich atomic nuclei. The fraction of free neutrons increases with increasing density, and at the bottom of the crust, at a density of about half nuclear saturation density, where nuclei occupy a significant fraction of space, nuclei are probably far from spherical. At this density nuclei disappear, and matter then becomes a uniform fluid of neutrons, protons and electrons. The thickness of this inner crust is typically about one kilometer. The EoS throughout the crust is sufficiently well understood for the purpose of building neutron star models.

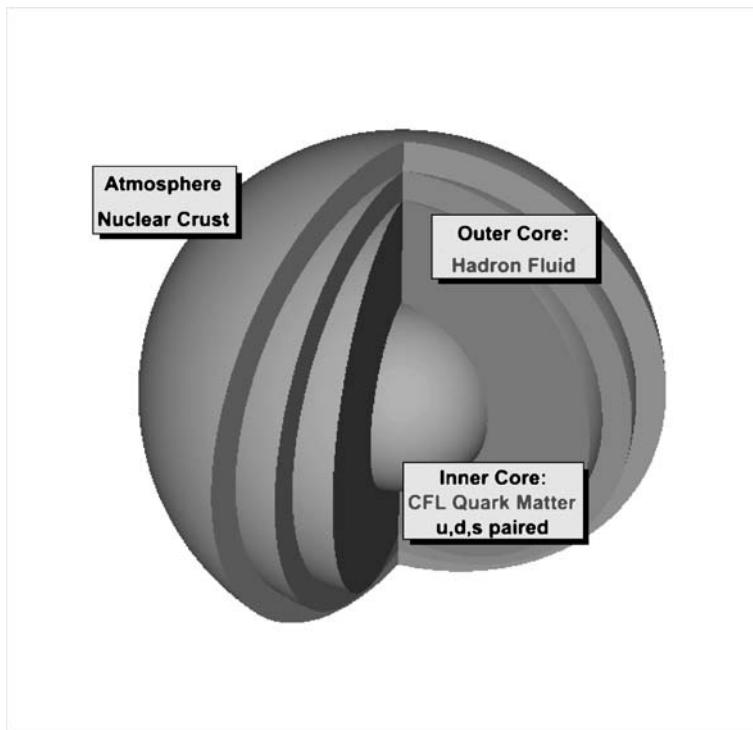


Fig. 6.25. Structure of neutron stars with masses $M > 1.2 M_{\odot}$ including a quark matter core

- Below the inner crust lies the **outer core**. At densities around nuclear density, matter consists of neutrons with a small admixture of protons, electrons and muons. All constituents are strongly degenerate. The neutrons and protons, which interact via nuclear forces, constitute a strongly **nonideal liquid**. The EoS is still not yet known with sufficient accuracy in the range above twice nuclear density.
- When the density reaches three times nuclear density, the individual bags of the nucleons overlap and the nuclear liquid makes a phase transformation to a quark bag. The exact value of the transition density is still uncertain, but most probably is in the range of two to four times nuclear density. Quark matter builds the **inner core** at least in stars more massive than about one solar mass. Quark matter at low temperature constitutes a color-superconducting fluid, where the quarks appear in Cooper pairs and the gluons become massive. Quark matter is however not a normal superconductor, since the photon remains massless. In distinction to nuclear matter, quark matter is asymptotically free, i.e. the EoS at high densities tends towards a free relativistic gas, $P \rightarrow \varrho/3$. As a consequence, neutron stars including quark cores are much more compact and their maximum mass is probably limited by $1.9 M_{\odot}$. The exact value depends on the interaction of the gluons with the quarks.

We can now ask what significance mass and radius measurements would have for the question of presence of quark matter in compact stars:

- An observed mass $M > 2 M_{\odot}$ would probably be inconsistent with the presence of quark matter in the core of the neutron star. By the inclusion of QCD corrections in terms of the parameter c , the maximum mass of originally 1.5 solar masses can be increased to about 1.9 solar masses for neutron stars including quark matter cores.
- Stars including a quark matter core are in general more compact than mean field models or pure nuclear cores (based on APR98 or SLy4 EoS). These stars have typically radii of only about 10 km at 1.4 solar masses.
- Could one conclude that color-superconducting matter is inside compact objects? This is a complicated question, since the actual value of the bag constant is a crucial parameter. Field theoretic models such as the NJL model (Fig. 6.19) predict the existence of the color-superconducting phases, but the details of the phase transitions are still uncertain.

6.4 Neutron Stars in Close Binary Systems

The masses of isolated neutron stars cannot be derived from observations. We have to find neutron stars in close binary orbits. The best candidates are then radio pulsars in closed orbit around a second compact object (white dwarf, neutron star or black hole). Exact masses can then be derived by observing post-Newtonian effects in close binary systems (periastron advance and gravitational redshift effects). In a few systems, even the Shapiro time-delay effect can now be observed.

6.4.1 Post-Newtonian Potentials for Many-Body Systems

The most accurate masses for neutron stars follow from the analysis of the motion of a radio pulsar in a close binary system. Since both masses are of the same order of magnitude, we have to treat a real relativistic two-body problem. Neutron stars represent extremely compact objects, they can be considered as point particles, at least for separations larger than 10 neutron star radii. The typical separation in a close binary system is about one solar radius, corresponding to at least 50,000 neutron star radii. On these scales, tidal effects can be completely neglected. As is well known from the analysis of planetary motion in the Solar System, relativistic gravity affects the motion in two ways. On the one hand, eccentric orbits undergo a perihelion shift, and, on the other hand, the orbits slowly shrink by the emission of gravitational waves. A third effect is that photons suffer a time delay when they propagate in the gravitational field of the Sun. All these effects will affect the propagation of the pulsar signal in a close binary system.

The Post-Newtonian Potentials

In order to analyze these effects in a close binary system, we have to consider the relativistic two-body system for two point masses. For this purpose we expand the metric tensor in post-Newtonian order, i.e. in the next higher order to the Newtonian expansion in the small parameter $\epsilon = V/c \simeq \sqrt{GM/c^2r}$, where r denotes the separation between the bodies. The expansion in the metric will be dictated, e.g. by considering the motion of a particle in the resulting gravitational field, following geodesics

$$\int_1^2 d\tau = \int_1^2 dt \sqrt{-g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}}. \quad (6.169)$$

Introducing the velocity $v^i = dx^i/dt$ and time $x^0 = t$, we arrive at the expression

$$\int_1^2 d\tau = \int_1^2 dt \sqrt{-(g_{tt} + 2g_{ti}v^i + g_{ik}v^iv^k)}. \quad (6.170)$$

The Newtonian expansion would be $g_{tt} \simeq -(1+2\Phi) = -1 + g_{tt}^{(2)}$, as well as $g_{ti} = 0$ and $g_{ik} = \delta_{ik}$. In the next order of the expansion we get the **post-Newtonian** form of the metric

$$g_{tt} = -1 + g_{tt}^{(2)} + g_{tt}^{(4)} + O(\epsilon^6) \quad (6.171)$$

$$g_{ti} = g_{ti}^{(3)} + O(\epsilon^5) \quad (6.172)$$

$$g_{ik} = \delta_{ik} + g_{ik}^{(2)} + O(\epsilon^4). \quad (6.173)$$

The superscripts denote the order of expansion in ϵ . Due to the equations of motion, the expansion in g_{tt} goes to order ϵ^{2n} , the expansion in g_{ti} to order ϵ^{2n-1} and the expansion in g_{ik} to order ϵ^{2n-2} . $n = 2$ is called post-Newtonian expansion, $n = 3$ post-post-Newtonian expansion, etc. The action for a test particle motion follows therefore from

$$\begin{aligned} \int ds &= \int dt \left[1 - \frac{1}{2}\mathbf{v}^2 + \frac{1}{2}g_{tt}^{(2)} \right] \\ &+ \int dt \left[-\frac{1}{8}\mathbf{v}^4 + \frac{1}{4}g_{tt}^{(2)}\mathbf{v}^2 + \frac{1}{2}g_{ik}^{(2)}v^iv^k + \frac{1}{2}g_{tt}^{(4)} - \frac{1}{8}\left(g_{tt}^{(2)}\right)^2 + g_{ti}^{(3)}v^i \right]. \end{aligned} \quad (6.174)$$

This corresponds to the formal expansion in the weak field limit $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Here we have the freedom to impose four gauge conditions, which are suitably chosen as

$$h_{tk,k} - \frac{1}{2}h_{kk,t} = 0 \quad (6.175)$$

$$h_{jk,k} + \frac{1}{2}h_{,j} = 0 \quad (6.176)$$

for the trace $h \equiv \eta_{\mu\nu}h^{\mu\nu} = -h_{tt} + h_{ii}$. These gauge conditions imply

$$g_{ik,k}^{(3)} - \frac{1}{2}g_{kk,t}^{(2)} = 0 \quad (6.177)$$

$$\frac{1}{2}g_{tt,i}^{(2)} + g_{ik,k}^{(2)} - \frac{1}{2}g_{kk,i}^{(2)} = 0. \quad (6.178)$$

The metric corrections h_{tt} , h_{ii} and h_{ik} should transform under spatial rotations as a scalar, vector and symmetric tensor, respectively, and thus should be constructed out of corresponding matter fields. For variables associated with the matter distribution we have the scalars ϱ_0 , $|\mathbf{x} - \mathbf{x}'|$, \mathbf{V}^2 , $\mathbf{V} \cdot (\mathbf{x} - \mathbf{x}')$, etc.; vector fields $\mathbf{x} - \mathbf{x}'$, \mathbf{V} , as well as tensors $\mathbf{V} \otimes \mathbf{V}$ and $(\mathbf{x} - \mathbf{x}') \otimes (\mathbf{x} - \mathbf{x}')$. Out of these quantities we can therefore construct the following potentials:

- a scalar potential

$$U(t, \mathbf{x}) = \int \frac{\varrho_0(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \quad (6.179)$$

- a tensor potential

$$U_{ik}(t, \mathbf{x}) = \int \frac{\varrho_0(t, \mathbf{x}') (x - x')^i (x - x')^k}{|\mathbf{x} - \mathbf{x}'|^3} d^3x', \quad (6.180)$$

- a kind of superpotential

$$\chi(t, \mathbf{x}) = - \int \varrho_0(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'| d^3x', \quad (6.181)$$

which satisfies

$$\chi_{,ik} = -\delta_{ik}U + U_{ik} \quad , \quad \nabla^2\chi = -2U, \quad (6.182)$$

- two velocity potentials (momentum currents)

$$V_i(t, \mathbf{x}) = \int \frac{\varrho_0(t, \mathbf{x}') v_i}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (6.183)$$

$$W_i(t, \mathbf{x}) = \int \frac{\varrho_0(t, \mathbf{x}') [\mathbf{v}' \cdot (\mathbf{x} - \mathbf{x}')] (x - x')_i}{|\mathbf{x} - \mathbf{x}'|^3} d^3x', \quad (6.184)$$

- as well as four additional post-Newtonian scalar potentials

$$\Phi_1 = \int \frac{\varrho_0' v'^2}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (6.185)$$

$$\Phi_2 = \int \frac{\varrho_0' U'}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (6.186)$$

$$\Phi_3 = \int \frac{\varrho_0' \Pi'}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (6.187)$$

$$\Phi_4 = \int \frac{P'}{|\mathbf{x} - \mathbf{x}'|} d^3x'. \quad (6.188)$$

The stress tensor is the source of these potentials

$$T^{tt} = \varrho_0(1 + \Pi + v^2 + U) \quad (6.189)$$

$$T^{ti} = \varrho_0(1 + \Pi + v^2 + 2U + P/\varrho_0)v^i \quad (6.190)$$

$$T^{ik} = \varrho_0 v^i v^k (1 + \Pi + v^2 + 2U + P/\varrho_0) + P\delta^{ik}(1 - 2\gamma U) \quad (6.191)$$

with Π as a measure for the internal energy.

In fact the solutions of any metric gravity theory can be expanded in terms of these potentials as follows

$$g_{tt} = -(1 - 2U + 2\beta U^2 - 4\Psi) \quad (6.192)$$

$$g_{ik} = (1 + 2\gamma U)\delta_{ik} \quad (6.193)$$

$$g_{ti} = -\frac{7}{2}\Delta_1 V_i - \frac{1}{2}\Delta_2 W_i \quad (6.194)$$

with the post-Newtonian parameters $\beta = 1$, $\gamma = 1$, $\Delta_1 = (4\gamma + 3)/7 = 1$ and $\Delta_2 = 1$ in general relativity. The post-Newtonian potential Ψ is given by a combination of the potentials Φ_A

$$\Psi = \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4. \quad (6.195)$$

In order to show the correctness of this ansatz for the metric, one has to solve Einstein's equations up to post-Newtonian order, i.e. to calculate the Christoffel symbols and Ricci tensors for the above expansion (which is left as an exercise).

For point particles of masses M_A these potentials reduce to the following form ($G = 1$)

$$U(\mathbf{r}_A) = \sum_{B \neq A} \frac{M_B}{r_{AB}} \quad (6.196)$$

$$\Phi_1(\mathbf{r}_A) = \sum_{B \neq A} \frac{M_B \mathbf{v}_B^2}{r_{AB}} \quad (6.197)$$

$$\Phi_2(\mathbf{r}_A) = \sum_{B \neq A} \frac{M_B}{r_{AB}} \sum_{C \neq B} \frac{M_C}{r_{BC}} \quad (6.198)$$

$$\Phi_3(\mathbf{r}_A) = 0 = \Phi_4(\mathbf{r}_A) \quad (6.199)$$

$$V_i(\mathbf{r}_A) = \sum_{B \neq A} \frac{M_B v^i}{r_{AB}} \quad (6.200)$$

$$W_i(\mathbf{r}_A) = \sum_{B \neq A} \frac{M_B (\mathbf{r}_{AB} \cdot \mathbf{v}_B) \mathbf{r}_{AB}}{r_{AB}^3} \quad (6.201)$$

The equation of motion for a particle could be derived directly from the geodesic equation, using

$$\begin{aligned} \frac{d^2 x^i}{dt^2} &= \frac{d}{dt} \left(\frac{dx^i}{dt} \frac{d\tau}{dt} \right) = \left(\frac{dt}{d\tau} \right)^{-1} \frac{d}{d\tau} \left[\left(\frac{dt}{d\tau} \right)^{-1} \frac{dx^i}{d\tau} \right] \\ &= \left(\frac{dt}{d\tau} \right)^{-2} \frac{d^2 x^i}{d\tau^2} - \left(\frac{dt}{d\tau} \right)^{-3} \frac{d^2 t}{d\tau^2} \frac{dx^i}{d\tau} \end{aligned} \quad (6.202)$$

ending up with the expression for the acceleration

$$\frac{d^2x^i}{dt^2} = -\Gamma_{\mu\nu}^i \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} + \Gamma_{\nu\lambda}^0 \frac{dx^\nu}{dt} \frac{dx^\lambda}{dt} \frac{dx^i}{dt}. \quad (6.203)$$

This can be expressed in terms of velocities

$$\frac{d^2x^i}{dt^2} = -\Gamma_{00}^i - 2\Gamma_{0k}^i v^k - \Gamma_{km}^i v^k v^m + \left[\Gamma_{00}^0 + 2\Gamma_{0j}^0 v^j + \Gamma_{jk}^0 v^j v^k \right] v^i. \quad (6.204)$$

The equations of motion can, however, be derived more easily by using the above action, or its Lagrangian for a point mass

$$\begin{aligned} \mathcal{L} = & -M_A \left(1 - \frac{1}{2} v_A^2 - \frac{1}{8} v_A^4 \right) \\ & + \sum_{A \neq B} \frac{M_A M_B}{r_{AB}} \left[1 + (2\gamma + 1)v_A^2 - (2\beta - 1) \sum_{C \neq A} \frac{M_C}{r_{AC}} - \frac{7}{2} \Delta_1 \mathbf{v}_A \cdot \mathbf{v}_B \right]. \end{aligned} \quad (6.205)$$

The total Lagrangian for an ensemble of particles with masses M_A can be written as (Einstein–Infeld–Hoffmann (EIH) Lagrangian for a N -body system)

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \sum_A M_A \mathbf{v}_A^2 + \frac{G}{2} \sum_{B \neq A} \frac{M_A M_B}{r_{AB}} \\ & + \frac{1}{8} \sum_A M_A \mathbf{v}_A^4 + \frac{3G}{2} \sum_A M_A \mathbf{v}_A^2 \sum_{B \neq A} \frac{M_B}{r_{AB}} \\ & - \frac{G}{4} \sum_{A \neq B} \frac{M_A M_B}{r_{AB}} [7\mathbf{v}_A \cdot \mathbf{v}_B + (\mathbf{v}_A \cdot \mathbf{n}_{AB})(\mathbf{v}_B \cdot \mathbf{n}_{AB})] \\ & - \frac{G^2}{2} \sum_A \sum_{B \neq C} \sum_{C \neq A} \frac{M_A M_B M_C}{r_{AB} r_{AC}}. \end{aligned} \quad (6.206)$$

Thereby, $\mathbf{n}_{AB} \equiv (\mathbf{x}_A - \mathbf{x}_B)/r_{AB}$ is the unit vector pointing in the direction of the two particles. This generalizes the Newtonian Lagrangian for a self-gravitating N -body system to post-Newtonian order. One can show that the equations of motion following from this Lagrangian are identical to (6.204), when the Christoffels are expressed in terms of the many-body potentials.

6.4.2 Periastron Shift in Two-Body Systems

For a two-body system this Lagrangian reduces to, $r \equiv r_{12}$ and $\mathbf{n} \equiv \mathbf{n}_{12}$,

$$\begin{aligned} \mathcal{L} = & \frac{M_1}{2} \mathbf{v}_1^2 + \frac{M_2}{2} \mathbf{v}_2^2 + \frac{1}{8} (M_1 \mathbf{v}_1^4 + M_2 \mathbf{v}_2^4) + \frac{GM_1 M_2}{r} \\ & + \frac{GM_1 M_2}{2r} [3(\mathbf{v}_1^2 + \mathbf{v}_2^2) - 7\mathbf{v}_1 \cdot \mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{n})(\mathbf{v}_2 \cdot \mathbf{n})] \\ & - \frac{G^2}{2} \frac{M_1 M_2 (M_1 + M_2)}{r^2}. \end{aligned} \quad (6.207)$$

One can define an effective mass

$$M_A^* = M_A \left(1 + \frac{1}{2} \mathbf{v}_A^2 \right) - \frac{1}{2} \frac{GM_A M_B}{r_{AB}} \quad (6.208)$$

and the EIH equations imply the conservation of the center-of-mass

$$\mathbf{X} \equiv \frac{M_1^* \mathbf{x}_1 + M_2^* \mathbf{x}_2}{M_1^* + M_2^*} \quad (6.209)$$

with

$$\frac{d^2 \mathbf{X}}{dt^2} = 0. \quad (6.210)$$

We now choose coordinates with $\mathbf{X} = 0$ and define reduced variables, $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$,

$$\mathbf{x}_1 = \left[\frac{M_2}{M} + \frac{\mu \delta M}{2M^2} \left(\mathbf{v}^2 - \frac{M}{r} \right) \right] \mathbf{x} \quad (6.211)$$

$$\mathbf{x}_2 = \left[\frac{M_1}{M} + \frac{\mu \delta M}{2M^2} \left(\mathbf{v}^2 - \frac{M}{r} \right) \right] \mathbf{x} \quad (6.212)$$

$$(6.213)$$

and get a reduced Lagrangian (modulo μ), $\mu = M_1 M_2 / M$, $\delta M = M_1 - M_2$,

$$\mathcal{L} = L_0 + L_1 \quad (6.214)$$

with

$$L_0 = \frac{1}{2} \mathbf{v}^2 + \frac{GM}{r} \quad (6.215)$$

$$L_1 = \frac{1}{8} \left(1 - \frac{3\mu}{M} \right) \mathbf{v}^4 + \frac{GM}{2r} \left[3\mathbf{v}^2 + \frac{\mu}{M} \mathbf{v}^2 + \frac{\mu}{M} (\mathbf{v} \cdot \mathbf{x}/r)^2 \right] - \frac{G^2 M^2}{2r^2}. \quad (6.216)$$

The corresponding Euler equation is then

$$\dot{\mathbf{v}} = -\frac{GM}{r^3} \mathbf{x} \left[1 - \frac{GM}{r} (4 + 2\mu/M) + (1 + 3\mu/M) \mathbf{v}^2 - (3\mu/2M) (\mathbf{v} \cdot \mathbf{x}/r)^2 \right] + \frac{GM}{r^3} \mathbf{v} (\mathbf{v} \cdot \mathbf{x}) (4 - 2\mu/M). \quad (6.217)$$

Similar to the classical orbital solution, the post-Newtonian solution for the orbit can be found in an analogous manner

$$\frac{p}{r} = 1 + e \cos \phi - \frac{M}{p} [(3 - \mu/M) + (1 + 9\mu/4M)e^2] \quad (6.218)$$

$$+ \frac{M}{p} \left[\frac{1}{2} (7 - 2\mu/M)e \cos \phi + 3e \phi \sin \phi - (\mu/4M)e^2 \cos 2\phi \right].$$

The term $\propto \phi$ represents a secular advance of the orbit with, $p = a(1 - e^2)$,

$$\boxed{\Delta\phi = \frac{6\pi G(M_1 + M_2)}{c^2 p}} \quad (6.219)$$

as the **periastron shift** per orbit. This expression is similar to the result obtained in Schwarzschild (see Sect. 8.1), except that the total mass $M = M_1 + M_2$ of the binary system is now involved³.

For the binary pulsar PSR 1913+16 the measured periastron shift is

$$\dot{\omega} = 4.226 \text{ deg yr}^{-1}. \quad (6.223)$$

The GR prediction is

$$\dot{\omega}_{GR} = 2.11 \left(\frac{M_1 + M_2}{M_\odot} \right)^{2/3} \text{ deg yr}^{-1}. \quad (6.224)$$

Due to the compactness of neutron stars, normal tidal effects are excluded. In this way one gets a handle on the **total mass of the system**

$$M_1 + M_2 = 2.85 M_\odot. \quad (6.225)$$

6.4.3 The Shapiro Time Delay in a Binary System

Signals propagating in the gravitational field of a binary system will suffer a time delay. This effect is usually discussed only for the Solar System, where it has been accurately measured. Photon propagation is given by null geodesics

$$g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} = 0. \quad (6.226)$$

³ There is a general way to derive any periastron shift in a perturbed orbit by using the Hamiltonian–Jacobi equation for a Hamiltonian of the form

$$H = H_0 + H_1. \quad (6.220)$$

H_0 is the Hamiltonian of the unperturbed problem

$$H_0 = \frac{\mathbf{p}_1^2}{2M_1} + \frac{\mathbf{p}_2^2}{2M_2} - \frac{GM_1M_2}{r}. \quad (6.221)$$

The perturbed Hamiltonian is just $H_1 = -L_1$, however expressed in terms of positions and momenta

$$\begin{aligned} H_1 = & -\frac{1}{8} \left(\frac{\mathbf{p}_1^4}{M_1^2} + \frac{\mathbf{p}_2^4}{M_2^2} \right) \\ & -\frac{G}{2r} \left[3 \left(\frac{M_2}{M_1} \mathbf{p}_1^2 + \frac{M_1}{M_2} \mathbf{p}_2^2 \right) - 7 \mathbf{p}_1 \cdot \mathbf{p}_2 - (\mathbf{p}_1 \cdot \mathbf{n})(\mathbf{p}_2 \cdot \mathbf{n}) \right] \\ & + \frac{G^2}{2} \frac{M_1 M_2 (M_1 + M_2)}{r^2}. \end{aligned} \quad (6.222)$$

The geodesic equation is given in post-Newtonian order

$$1 - 2U - \left| \frac{d\mathbf{x}}{dt} \right|^2 (1 + 2\gamma U) = 0. \quad (6.227)$$

The resulting trajectory of a signal in the binary system is in first approximation a straight line with some small deviation, \mathbf{x}_{PSR} denoting the position of the pulsar,

$$\mathbf{x}(t) = \mathbf{x}_{\text{PSR}} + \mathbf{n}(t - t_e) + \mathbf{x}_p(t), \quad (6.228)$$

where we assumed $\mathbf{x}_p(t_e) = 0$. The signal travels initially in the direction of the unit vector \mathbf{n} and is then deflected by the gravitational field of the companion. It is useful to decompose \mathbf{x}_p into components parallel and perpendicular to the unperturbed trajectory

$$x_p(t)_{\parallel} = \mathbf{n} \cdot \mathbf{x}_p \quad (6.229)$$

$$x_p(t)_{\perp} = \mathbf{x}_p(t) - \mathbf{n}[\mathbf{n} \cdot \mathbf{x}_p]. \quad (6.230)$$

From the null geodesics condition we get

$$\frac{dx_{p\parallel}}{dt} = -(1 + \gamma)U \quad (6.231)$$

and from the geodesics equation we can calculate the acceleration

$$\frac{d^2x_{p\perp}^i}{dt^2} = (1 + \gamma)[U_{,i} - \mathbf{n}^i(\mathbf{n} \cdot \nabla U)]. \quad (6.232)$$

Along the unperturbed path of the signal, U has the form

$$U = \frac{GM_2}{r(t)} = \frac{GM_2}{|\mathbf{x}_P + \mathbf{n}(t - t_e)|}. \quad (6.233)$$

The equation for $x_p(t)$ can easily be integrated to yield

$$x_{p\parallel}(t) = -(1 + \gamma) \frac{GM_2}{c^2} \ln \left(\frac{r(t) + \mathbf{n} \cdot \mathbf{x}(t)}{r_P + \mathbf{n} \cdot \mathbf{x}_P} \right),$$

(6.234)

which obviously satisfies the initial condition $x_{p\parallel}(t_e) = 0$.

6.4.4 Decay of Binary Orbits due to Gravitational Radiation

In contrast to electrodynamics, gravitational radiation is generated by time-dependent mass quadrupole moments $Q_{ik}(t)$ with an energy loss first derived by Einstein in 1917

$$\frac{dE}{dt} = -\frac{G}{45c^5} \ddot{\mathcal{Q}}_{ik} \dot{\mathcal{Q}}_{ik}.$$

(6.235)

The third time derivative of the traceless mass quadrupole tensor is the source of the energy loss

$$Q_{ik} = \int (3x^i x^k - \delta_{ik} \mathbf{x}^2) \varrho_0(t, \mathbf{x}) d^3 x. \quad (6.236)$$

For the calculation one needs the solutions of a binary system (in units $G = 1 = c$) in terms of the conserved angular momentum L and total energy $E < 0$

$$a = -\frac{M_1 M_2}{2E}, \quad E = -\frac{1}{2} \frac{M_1 M_2}{a} \quad (6.237)$$

$$e^2 = 1 + \frac{2EL^2(M_1 + M_2)}{M_1^3 M_2^3} \quad (6.238)$$

$$r = \frac{a(1 - e^2)}{1 + e \cos \phi} \quad (6.239)$$

$$r_1 = \frac{M_2}{M_1 + M_2} r \quad (6.240)$$

$$r_2 = \frac{M_1}{M_1 + M_2} r. \quad (6.241)$$

We first compute the moments of inertia

$$I_{ik} = \int \varrho_0(\mathbf{x}) x^i x^k d^3 x \quad (6.242)$$

and choose the orbital plane given by coordinates (x, y) . This produces the following components

$$I_{xx} = M_1 x_1^2 + M_2 x_2^2 = \frac{M_1 M_2}{M_1 + M_2} r^2 \cos^2 \phi \quad (6.243)$$

$$I_{yy} = \frac{M_1 M_2}{M_1 + M_2} r^2 \sin^2 \phi \quad (6.244)$$

$$I_{xy} = \frac{M_1 M_2}{M_1 + M_2} r^2 \sin \phi \cos \phi \quad (6.245)$$

$$I = I_{xx} + I_{yy} = \frac{M_1 M_2}{M_1 + M_2} r^2. \quad (6.246)$$

These quantities determine the energy loss

$$-\frac{dE}{dt} = \frac{1}{5} \left[\ddot{I}_{ik} \ddot{I}_{ik} - \frac{1}{3} (\ddot{I})^2 \right] = \frac{1}{5} \left[(\ddot{I}_{xx})^2 + (\ddot{I}_{yy})^2 + 2(\ddot{I}_{xy})^2 - \frac{1}{3} (\ddot{I})^2 \right]. \quad (6.247)$$

Inserting these expressions into the energy-loss formula we find (see problem 6.6)

$$-\frac{dE}{dt} = \frac{8M_1^2 M_2^2}{15a^2(1 - e^2)^2} \left[12(1 + e \cos \phi)^2 + e^2 \sin^2 \phi \right] \dot{\phi}. \quad (6.248)$$

This loss is averaged over one orbital period P_b according to Kepler's law

$$P_b = \frac{2\pi a^{3/2}}{\sqrt{M_1 + M_2}} \quad (6.249)$$

with

$$-\frac{dE}{dt} = -\frac{1}{P_b} \int_0^{P_b} \frac{dE}{dt} dt = -\frac{1}{P_b} \int_0^{2\pi} \frac{dE}{dt} \frac{1}{\dot{\phi}} d\phi. \quad (6.250)$$

This integral can be done explicitly

$$-\frac{dE}{dt} = \frac{32}{5} \frac{M_1^2 M_2^2 (M_1 + M_2)}{a^5 (1 - e^2)^{7/2}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right), \quad (6.251)$$

or in dimensional form (Wagoner [408])

$$-\frac{dE}{dt} = \frac{32}{5} \frac{G^4 M_1^2 M_2^2 (M_1 + M_2)}{a^5 c^5 (1 - e^2)^{7/2}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right). \quad (6.252)$$

This has in fact the correct dimension for an energy loss. For the binary system PSR B1913+16, the energy loss is considerable

$$-\frac{dE}{dt} = 0.6 \times 10^{33} \text{ erg s}^{-1} = 0.16 L_\odot. \quad (6.253)$$

The change in the orbital period follows from

$$\frac{1}{P_b} \frac{dP_b}{dt} = -\frac{3}{2} \frac{1}{E} \frac{dE}{dt} \quad (6.254)$$

and hence for PSR B1913+16 $\dot{P}_b = -2.4 \times 10^{-12} \text{ s s}^{-1}$. The shrinking of the orbit follows from

$$\begin{aligned} \frac{da}{dt} &= \frac{2a^2}{GM_1 M_2} \frac{dE}{dt} \\ &= -\frac{64}{5} \frac{G^3 M_1 M_2 (M_1 + M_2)}{a^3 c^5 (1 - e^2)^{7/2}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right). \end{aligned} \quad (6.255)$$

The solution to this equation can be written as, $a_0 = a(0)$,

$$a(t) = a_0 \left(1 - \frac{t}{t_\infty} \right)^{1/4} \quad (6.256)$$

with the characteristic evolution time t_∞ (**merger time**) given by

$$t_\infty = \frac{5}{256} \frac{a_0^4 c^5 (1 - e^2)^{7/2}}{G^3 M_1 M_2 (M_1 + M_2)} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right)^{-1}. \quad (6.257)$$

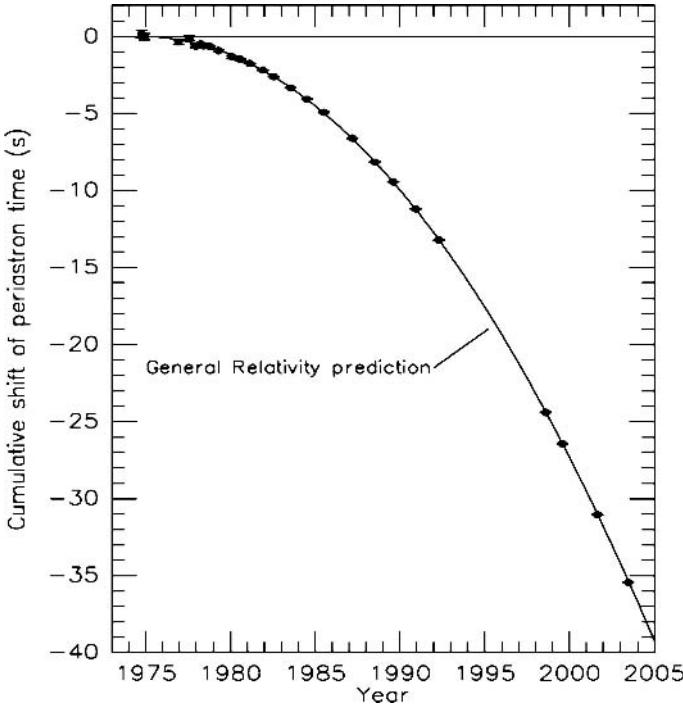


Fig. 6.26. Orbital decay of the binary system PSR 1913+16. The *data points* indicate the observed change in the epoch of periastron. The parabola corresponds to the theoretically expected change in the epoch for a system emitting gravitational waves according to general relativity. Figure adapted from Weisberg and Taylor [415]

This means that the characteristic time-scale is given by the light travel time through the orbit

$$t_\infty \simeq 0.02 \left(\frac{c^2 a_0}{G M_1} \right)^3 \frac{a_0}{c}, \quad (6.258)$$

prolonged by the third power in the inverse compactness of the orbit, $G M_1 / c^2 a_0 \simeq 10^{-6}$, giving a characteristic time-scale $t_\infty \simeq 10^{16} a_0 / c \simeq 10^{16}$ s \simeq a few hundred million years. These systems will merge into a single object on time-scales shorter than the age of the Universe. For this reason, such systems are thought to be possible candidates for the origin of gamma-ray bursters and intense gravitational wave emission.

In this discussion we have assumed that the eccentricity remains constant. Gravitational wave emission leads however also to a change in the eccentricity

$$e \frac{de}{dt} = \frac{M_1 + M_2}{M_1^3 M_2^3} \left(L^2 \frac{dE}{dt} + 2EL \frac{dL}{dt} \right), \quad (6.259)$$

which can be written as

$$\frac{1}{e} \frac{de}{dt} \frac{P_b dt}{dP_b} = \frac{19}{18} (1 - e^2) \left(1 + \frac{121}{304} e^2 \right) \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right)^{-1} = 0.34. \quad (6.260)$$

As a consequence, the eccentricity decreases with time. The reason for this is that gravitational waves carry away angular momentum. The change in the angular momentum can be written

$$\frac{dL}{dt} = -\frac{32}{5} \frac{G^{7/2} M_1^2 M_2^2 (M_1 + M_2)^{1/2}}{c^5 a^{7/2}} \left(1 + \frac{7}{2} e^2 \right) (1 - e^2)^{-1}. \quad (6.261)$$

6.5 Masses of Neutron Stars from Radio Pulsar Timing

After several pulsar surveys, there are now about 50 known radio pulsars in binary systems, including at least seven in double neutron star binaries (Thorsett and Chakrabarty [393]).

In some cases, the stellar masses can be directly determined from measurements of relativistic orbital effects. In others, only an indirect or statistical estimate of the masses is possible. In the following the general problem of mass measurement in radio pulsar binaries and all current estimates of the masses of radio pulsars and their companions are discussed. Significant constraints exist on the masses of 21 radio pulsars and on six neutron star companions of radio pulsars. All the measurements are consistent with a remarkably narrow underlying Gaussian mass distribution, $M = 1.35 \pm 0.04 M_\odot$. There is no evidence that extensive mass accretion ($\Delta M \geq 0.1 M_\odot$) has occurred in these systems. The observed inclinations of millisecond pulsar binaries are consistent with a random distribution, and thus there is no evidence for either alignment or counter-alignment of millisecond pulsar magnetic fields.

6.5.1 What is Pulsar Timing?

The orbiting pulsar is an ideal clock with an accuracy of $\dot{P} \simeq 10^{-18}$ for millisecond pulsars. The individual signals emitted by the pulsar are however somewhat washed out, so that the timing accuracy is only in the range of $10 \mu s \simeq 10^{-3} P$. This corresponds to about the travel time across the polar cap region of a neutron star. Since this has to do with the physical origin of the radio emission, it will be barely possible in the future to improve the timing accuracy. If we plot the pulse number N as a function of time, this would be a linear relation for an ideal clock. The deceleration of the pulsar, measured as $\dot{P} > 0$ makes a deviation from this linear relation. If the pulsar orbits in a binary system, this curve $N(t)$ will be modulated by means of the periodic orbit, i.e. the curve $N(t)$ wobbles periodically around the linear trend with an amplitude given by the light travel time across the binary system.

For a close binary system, this amplitude is of the order of seconds, and the period is given by the orbital period P_b .

Soon after the discovery of the system PSR 1913+16 Blandford and Teukolsky (1976) derived a timing model for fitting to a sequence of pulse arrival times on Earth. They assumed that the pulsar and its companion obeyed the Keplerian laws. In this model the following Keplerian parameters enter into the timing model: the period P_b , the epoch of periastron passage T_0 , the eccentricity e , the longitude of periastron ω and the projected semimajor axis $x = a_1 \sin i/c$, where a_1 is the semimajor axis of the pulsar orbit and i the angle of inclination (see Fig. 6.27). They also incorporated the largest relativistic effects, a combination of gravitational redshift and time dilation, quantified in a parameter called γ_{RD} . Epstein [144] and Haugen [196] completed the Blandford–Teukolsky model by deriving within GR the order $O(v^2/c^2)$ contributions to the timing formula. These terms arise from the Shapiro time delay in the gravitational potential of the companion and from post-Newtonian effects in the orbital motion. This formulation was given within the framework of GR. Damour and Deruelle [128] found a way to incorporate all

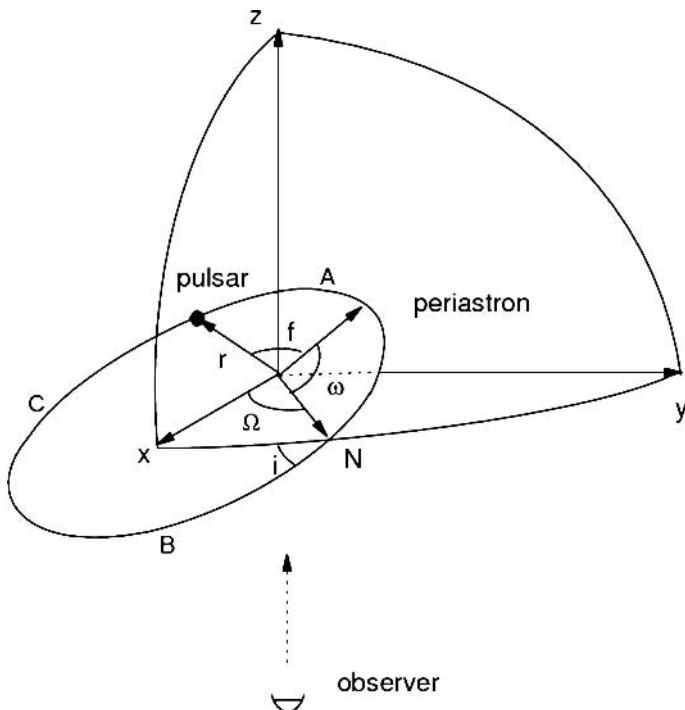


Fig. 6.27. Geometry and orbital elements for a pulsar in a binary system. In a first approximation, the orbits are still Keplerian ellipses, which however advance in prograde direction. Ω is the longitude of the ascending node N , ω the longitude of the periastron. i is the orbital inclination angle, and f the true anomaly

$O(v^2/c^2)$ timing effects in a simple mathematical way common to a large class of gravitational theories (see also Damour and Taylor [129]).

We now consider an inertial system in the center-of-mass of the binary system. In this coordinate system, an observer is given by (t_B, \mathbf{x}_B) , and the pulsar in terms of (t_P, \mathbf{x}_P) . The time difference is then

$$c(t_B - t_P) = |\mathbf{x}_B - \mathbf{x}_P| + \frac{2GM_2}{c^2} \ln \left(\frac{2d(t_B)}{r(t) + \mathbf{x}_P \cdot \mathbf{n}} \right) + c \Delta_{DM}. \quad (6.262)$$

The second term corresponds to the Shapiro time delay with an amplitude given by $\Delta_S \simeq R_S/c \simeq 15 \mu\text{s}$. The first factor can be expanded

$$|\mathbf{x}_B - \mathbf{x}_P| = d_B(t_B) - \mathbf{x}_1(t_P) \cdot \mathbf{n} + O(r/d_B), \quad (6.263)$$

where d_B denotes the distance of the observer with respect to the barycenter. From this we obtain

$$c(t_B - t_P) = -d_B + \mathbf{x}_1(t_P) \cdot \mathbf{n} = -d_B + \mathbf{x}_1(t_\infty - d_B/c + \mathbf{x}_1 \cdot \mathbf{n}/c) \cdot \mathbf{n} \quad (6.264)$$

and therefore in a Taylor expansion

$$c(t_B - t_P) \simeq -d_B + \mathbf{x}_1(t_\infty - d_B/c) \cdot \mathbf{n} + [\dot{\mathbf{x}}_1(t_\infty - d_B/c) \cdot \mathbf{n}] [\mathbf{x}_1(t_\infty - d_B/c) \cdot \mathbf{n}]. \quad (6.265)$$

The various terms have the following interpretation

- $\mathbf{x}_1 \cdot \mathbf{n} \simeq a_P \sin i/c \simeq 1 \text{s}$ is the **linear Doppler effect**;
- $\mathbf{x}_1 \cdot \mathbf{n} (v_1/c) \simeq 5 \text{ ms}$ is the **quadratic Doppler effect**;
- the next higher order term $\mathbf{x}_1 \cdot \mathbf{n} (v_1/c)^2 \simeq 10 \mu\text{s}$ includes **post-Newtonian corrections**, since the orbit is not exactly elliptic;
- the next higher order terms are no longer measurable, since they would have amplitudes in the range of 20 ns. This also corresponds to the accuracy for ephemerid calculations in the Solar System.

In the arrival time curve $N(t)$ we get a superposition of modulations of smaller and smaller amplitudes.

There is an additional effect: the post-Newtonian time t_P of the pulsar has to be transformed to its proper time T_P , corresponding to its clock. The line element provides the following relation between the proper time T_P of the radio pulsar and the post-Newtonian time t_P

$$dT_P^2 = -ds^2/c^2 = (1 - 2U) dt^2 - \mathbf{v}_P^2/c^2 dt^2, \quad (6.266)$$

with the following potential

$$U(t, \mathbf{x}) = \frac{GM_1}{c^2 |\mathbf{x} - \mathbf{x}_1(t)|} + \frac{GM_2}{c^2 |\mathbf{x} - \mathbf{x}_2(t)|}. \quad (6.267)$$

From this we get

$$\frac{dT_P}{dt} = \sqrt{1 - 2U - \mathbf{v}_P^2} \simeq 1 - U(t, \mathbf{x}_P) - \frac{1}{2}\mathbf{v}_P^2. \quad (6.268)$$

With the velocity v_p for a Newtonian orbit of semimajor axis a , $M = M_1 + M_2$ and $m_1 = GM_1/c^2$,

$$v_p^2 = \frac{M_2^2}{M} \left(\frac{2}{r} - \frac{1}{a} \right) \quad (6.269)$$

one gets

$$\begin{aligned} \frac{dT_P}{dt} &= 1 - \frac{M_2}{r} - \frac{M_2^2}{M} \frac{1}{r} + \text{const} \\ &= 1 - \frac{1}{r} \frac{M_2(M_1 + 2M_2)}{M} + \text{const}. \end{aligned} \quad (6.270)$$

The relative separation is given by

$$r = |\mathbf{x}| = a(1 - e \cos E). \quad (6.271)$$

Time is related to the eccentric anomaly E by means of Kepler's equation

$$\frac{2\pi}{P_b}(t_P - T_0) = E - e \sin E, \quad (6.272)$$

or

$$\frac{2\pi}{P_b} dt = (1 - e \cos E) dE \quad (6.273)$$

Therefore

$$dT_P = dt_P - \frac{M_2}{a} \left(1 + \frac{M_2}{M} \right) dE \quad (6.274)$$

and therefore by using once again the Kepler equation we find

$$T_P = t_P - \gamma_{RD} \sin E(t) - \frac{GM_2}{c^2 a} \left(1 + \frac{M_2}{M} \right) t + \text{const} \quad (6.275)$$

with

$$\gamma_{RD} = \frac{GM_2(M_1 + 2M_2)}{c^2 a M} \frac{e P_b}{2\pi}. \quad (6.276)$$

The third factor is suppressed with $(v_p/c)^2 \simeq 10^{-6}$. a can be replaced in terms of Kepler's third law so that

$\gamma_{RD} = e \left(\frac{P_b}{2\pi} \right)^{1/3} M_2(M_p + 2M_2) M^{-4/3}.$

(6.277)

In this form the effect is usually expressed in the literature. For circular orbits, the gravitational redshift is constant, and similarly the quadratic Doppler shift is also constant so that γ_{RD} vanishes. This effect is therefore important in highly eccentric binary systems.

6.5.2 The Timing Formula

Ideally, in order to model the rotational behavior of the neutron star, we require TOAs (times of arrival) measured by an inertial observer. Due to the Earth's orbit around the Sun, an observatory located on Earth experiences accelerations with respect to the neutron star. The observatory is therefore not in an inertial frame. To a very good approximation, the center-of-mass of the Solar System, the Solar System barycenter, can be regarded as an inertial frame. It is standard practice to transform the observed TOAs to this frame using a planetary ephemeris such as the JPL DE200. The transformation is summarized in the following equation as the difference between the time T in the comoving pulsar frame and the topocentric time t of the observer

$$T - t = t_0 + \Delta_C - D/v^2 + \Delta_{R\odot} + \Delta_{E\odot} + \Delta_{S\odot} - \Delta_R - \Delta_E - \Delta_S. \quad (6.278)$$

Here t_0 is a reference epoch and Δ_C is the offset between the observatory master clock and the reference standard of terrestrial time. The dispersive delay is proportional to $1/v^2$, where $D = \text{DM}/2.41 \times 10^{-4}$ with the dispersion measure DM in units of pc cm $^{-3}$, when the radio frequency v is in units of MHz and the delay is in seconds. This effect is important when TOAs for signals of different frequencies are compared. $\Delta_{R\odot}$, $\Delta_{E\odot}$ and $\Delta_{S\odot}$ are propagation delays and relativistic time adjustments for effects in the Solar System. Δ_R , Δ_E and Δ_S are the corresponding effects accounting for phenomena within the pulsar's orbit.

The first effect in a binary system is related to the varying propagation time of a signal, called Römer time delay Δ_R . The amplitude follows from the orbital motion of the pulsar in the binary system when projected into the orbital plane

$$\Delta_R = x \sin \omega (\cos E - e) + x \sqrt{1 - e^2} \cos \omega \sin E, \quad (6.279)$$

the next effect is due to varying gravitational redshift and quadratic Doppler effect

$$\Delta_E = \gamma_{\text{RD}} \sin E, \quad (6.280)$$

and the third effect results from the Shapiro time delay written as

$$\Delta_S = 2r \ln \left(1 - e \cos E - s [\sin \omega (\cos E - e) + \sqrt{1 - e^2} \cos \omega \sin E] \right). \quad (6.281)$$

r is called the range parameter and s the shape parameter. These terms are written in the eccentric anomaly E as a generalization of the Newtonian E , eccentricity e , the projected semimajor axis x in units of seconds, the longitude of periastron ω , and the epoch of periastron passage T_0 . The eccentric anomaly E is defined by the relation

$$E - e \sin E = 2\pi \left[\frac{T - T_0}{P_b} - \frac{\dot{P}_b}{2} \left(\frac{T - T_0}{P_b} \right)^2 \right], \quad (6.282)$$

where the second term includes a possible shrinking of the orbit by a change in the orbital period P_b . The factor $-1/2$ comes from the integration of the instantaneous orbital frequency $1/[P_b + \dot{P}_b(T - T_0)]$ to obtain the orbital phase. Similarly, the true anomaly $A(E)$ follows from

$$A(E) = 2 \arctan \left[\sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2} \right]. \quad (6.283)$$

The advance of periastron is then measured by

$$\omega = \omega_0 + \frac{P_b \dot{\omega}}{2\pi} A(E). \quad (6.284)$$

At a given time, the propagation delay across the pulsar's orbit is calculated by a model that contains 10 parameters

$$t - t_0 = F[T; P_b, e_0, T_0, \omega_0, x_0; \dot{\omega}, \gamma_{\text{RD}}, \dot{P}_b, r, s]. \quad (6.285)$$

This is called the direct **timing formula**. The contributions from the individual effects are shown in Fig. 6.29. When analyzing pulsar data one needs the inverse timing formula $T = f(t)$ in which the Solar System barycentric time t is the independent variable. The fit parameters are the five Keplerian parameters (x, ω, T_0, P_b, e) evaluated at epoch t_0 , and the five post-Keplerian parameters $(\dot{\omega}, \dot{P}_b, \gamma_{\text{RD}}, r, s)$. In a particular theory of gravity, the post-Keplerian parameters can be written as a function of the pulsar mass M_1 and of its companion M_2 , and the total mass $M = M_1 + M_2$ of the binary system. In general relativity one finds the following expressions

$$\dot{\omega} = 3 T_{\odot}^{2/3} \left(\frac{P_b}{2\pi} \right)^{-5/3} \frac{(M_1 + M_2)^{2/3}}{(1 - e^2)} \quad (6.286)$$

$$\gamma_{\text{RD}} = e T_{\odot}^{2/3} \left(\frac{P_b}{2\pi} \right)^{1/3} \frac{M_2(M_1 + 2M_2)}{M^{4/3}} \quad (6.287)$$

$$r = T_{\odot} M_2 \quad (6.288)$$

$$s = x \left(\frac{P_b}{2\pi} \right)^{-2/3} T_{\odot}^{-1/3} M^{2/3} M_2^{-1} \quad (6.289)$$

$$\dot{P}_b = -\frac{192\pi}{5} \left(\frac{2\pi T_{\odot}}{P_b} \right)^{5/3} \frac{1 + \frac{73}{24}e^2 + \frac{37}{96}e^4}{(1 - e^2)^{7/2}} \frac{M_1 M_2}{(M_1 + M_2)^{1/3}}. \quad (6.290)$$

The masses M_1 , M_2 and M are expressed in solar masses, $s \equiv \sin i$, and $T_{\odot} \equiv GM_{\odot}/c^3 = 4.925490947 \mu\text{s}$ is the unit of time known from Solar System timing.

Following the accumulation of about 10 to 20 barycentric TOAs from observations spaced over at least several months, a surprisingly simple model can be applied to the TOAs and optimized so that it is sufficient to account for the arrival time of any pulse emitted during the time span of the observations and predict the arrival times of subsequent pulses. The model is based on a Taylor expansion of the angular

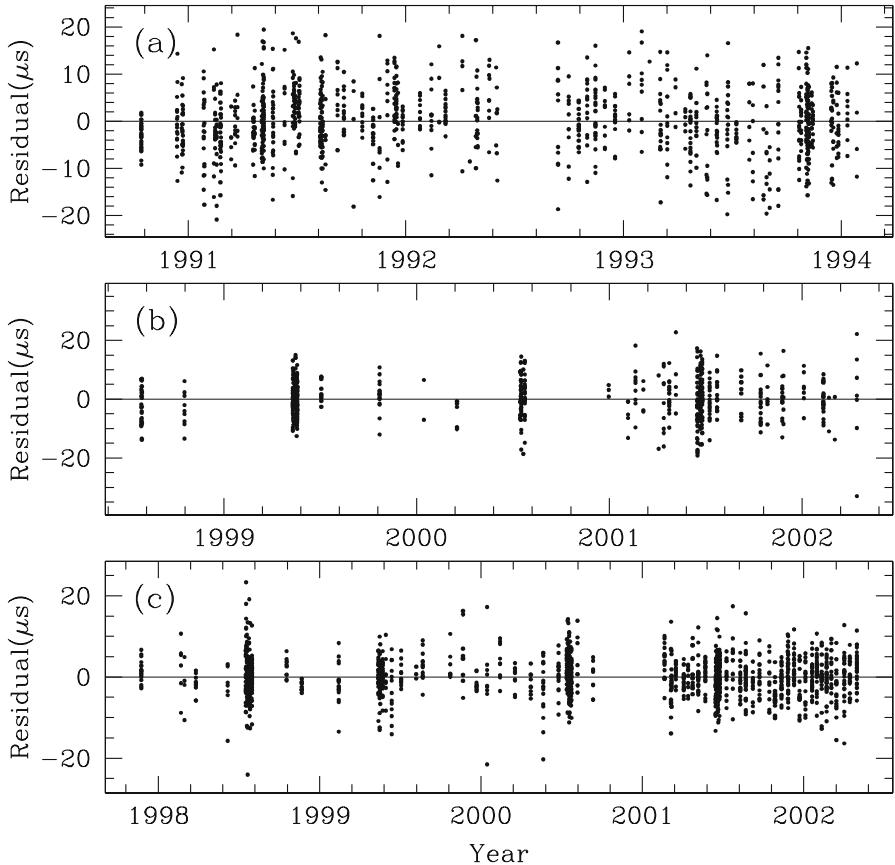


Fig. 6.28. Timing residuals for PSR B1534+12 remained constant over a decade. Figure adapted from [373]

rotational frequency about a model value at some reference epoch. The model pulse phase as a function of barycentric time is thus given by:

$$\Phi(T) = \Phi_* + (T - T_*)\Omega_* + \frac{1}{2}(T - T_0)^2\dot{\Omega}_* + \dots \quad (6.291)$$

where Φ_* is the pulse phase at T_* . Based on this simple model, and using initial estimates of the position, dispersion measure and pulse period, a *timing residual* is calculated for each TOA as the difference between the observed and predicted pulse phases (Fig. 6.28).

Early sets of residuals will exhibit a number of trends indicating a systematic error in one or more of the model parameters, or a parameter not initially incorporated into the model. An error in the assumed parameter set results in a linear slope with time. A parabolic trend results from an error. Additional effects will arise if the assumed position of the pulsar is incorrect. A position error of just one arcsecond

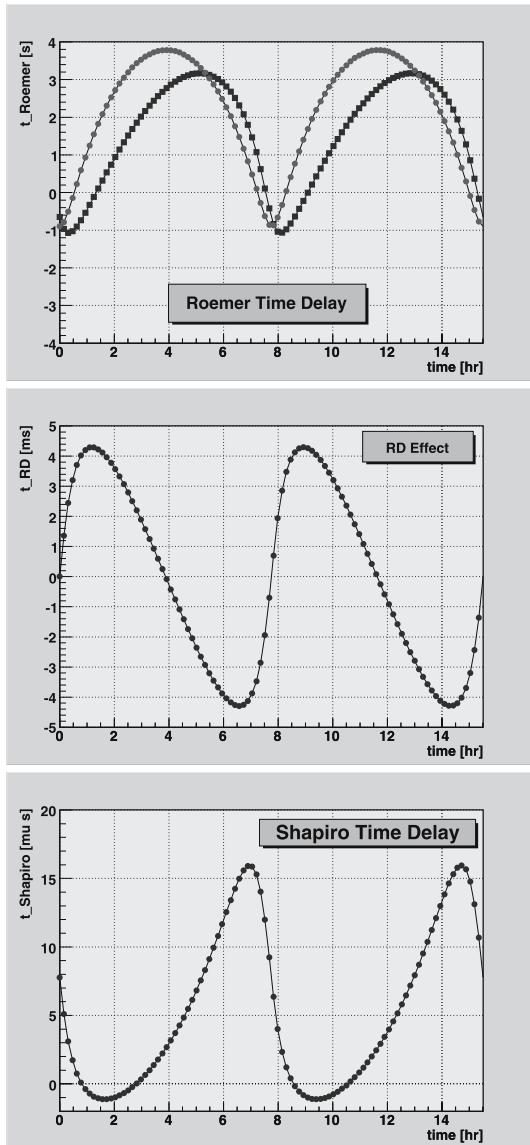


Fig. 6.29. Simulation of the modulation of pulsar arrival times as a function of the orbital time for the binary system PSR J1913+16. *Top:* modulation by means of the Roemer propagation time for two different epochs; *middle:* gravitational redshift and quadratic Doppler effect; *bottom:* Shapiro time delay due to the gravitational field of the companion. Since the amplitude of the Shapiro effect is of the same order as the timing accuracy, this effect cannot be observed in this system

results in an annual sinusoid with a peak-to-peak amplitude of about 5 ms for a pulsar on the ecliptic; this is easily measurable for typical TOA uncertainties of order one milliperiod or better. Similarly, the effect of a proper motion produces an annual sinusoid of linearly increasing magnitude. This procedure provides therefore also astrometric information (celestial coordinates and proper motion) of pulsars.

After a number of iterations, and with the benefit of a modicum of experience, it is possible to identify and account for each of these various effects to produce

Table 6.5. Timing data for the two most famous relativistic binary systems. The numbers in brackets give the errors. In the first group, you find the astrometric parameters, the second group gives the classical orbital elements of binary systems, and the third group lists the post-Keplerian parameters

Parameter	PSR B1913+16	PSR B1534+12
Spin period P [ms]	59.029997929613(7)	37.90444048785528(5)
Braking rate \dot{P} [10^{-18}]	8.62713(8)	2.42253(3)
Right ascension (J2000)	19:15:28.0002	15:37:09.95994(2)
Declination (J2000)	16:06:27.4043	11:55:55.6561(3)
Dispersion [pc cm^{-3}]	168.770	11.619(12)
Timing accuracy [μs]	15	3
Orbital period P_b [d]	0.322997462	0.42073729933(3)
Eccentricity e	0.6171308(4)	0.2736775(5)
Semimajor axis $a_p \sin i/c$ [s]	2.3417592(19)	3.729464(3)
Periastron length ω_0 [deg]	226.57528(6)	267.44746(16)
Periastron passage T_0 [MJD]	46443.99588319(3)	48778.82595096
Periastron shift $\dot{\omega}$ [deg/yr]	4.226621(11)	1.755794(19)
Grav/Doppler effect γ_{RD} [ms]	4.295(2)	2.071(6)
Shapiro time delay r [μs]	–	6.7(1,3)
Orbital inclination $s = \sin i$	–	0.983(8)
Orbit decay \dot{P}_b [10^{-12} s/s]	-2.422(6)	-0.129(14)

a timing solution which is phase coherent over the whole data span. The resulting model parameters provide spin and astrometric information about the neutron star to a precision which improves as the length of the data span increases (see Table 6.5). The latest observations of the original millisecond pulsar, B1937+21, spanning almost 9 yr (exactly 165,711,423,279 rotations!) measure a period of $1.5578064688197945 \pm 0.0000000000000004$ ms defined at midnight UT on December 5 1988! Astrometric measurements based on these data are no less impressive, with position errors of arcsec being presently possible.

6.5.3 Timing of the Binary System PSR B1913+16

More than seven binary systems consisting of a pulsating neutron star have been found (Table 6.6). The merger time determined by emission of gravitational waves is given by

$$\tau_{\text{merge}} \simeq 10^7 \text{ yr} \left(\frac{P_b}{10 \text{ hr}} \right)^{8/3} \left(\frac{\mu}{M} \right)^{-1} \left(\frac{M}{M_\odot} \right)^{-2/3} (1 - e^2)^{7/2}. \quad (6.292)$$

The solution of the timing procedure is given in Table 6.5 for the the two systems PSR B1913+16 and PSR B1534+12. The Shapiro parameters are not yet well defined for PSR B1913+16 due to its low inclination. The post-Keplerian parameters can now

Table 6.6. Binary systems containing radio pulsars which coalesce in less than a Hubble time. τ_P is the pulsar characteristic age, τ_{merge} is the time remaining to coalesce due to emission of gravitational waves. B2127+11C is in a globular cluster

Pulsar	P [ms]	P_b [hr]	e	Mass [M_\odot]	τ_P [Myr]	τ_{merge} [Myr]	Detection
J0737–3039A	22.70	2.45	0.088	2.58	210	87	2003
J0737–3039B	2773	2.45	0.088	2.58	50	87	2004
B1534+12	37.90	10.10	0.274	2.75	248	2690	1990
J1756–2251	28.46	7.67	9.181	2.57	444	1690	2004
B1913+16	59.03	7.75	0.617	2.83	108	310	1975
B2127+11C	30.53	8.04	0.681	2.71	969	220	1990
J1141–6545	393.90	4.74	0.172	2.30	1.4	590	2000
J1518+4904	40.9	8.63d	0.249	2.62		9600	
J1811–1736	104.2	18.78d	0.828	2.60		1700	
J1829+2456	41.0	1.17d	0.14	2.53		60	

be used to constrain the position of the pulsar in the mass plot (M_1, M_2) (Fig. 6.30). The crossing of the lines given by the periastron shift $\dot{\omega}$ and the quadratic Doppler effect γ_{RD} determine the position very accurately. The Shapiro shape parameter which determines the inclination $\sin i$ of the orbital plane gives a first consistency check.

A further relation is obtained from the shrinking of the orbital motion by the emission of gravitational waves.

6.5.4 Masses of Companion Stars

While timing measurements of the relativistic corrections to the Keplerian orbital equations provide the most accurate mass determinations for neutron stars, these methods are only possible for close eccentric binary orbits or when the orbit is viewed more or less edge-on. In the great majority of binary systems, the mass function provides the only timing information about the companion mass. The pulsar mass can then only be determined if additional constraints are found for the companion star.

In recent years, about a dozen companion stars in radio pulsar binaries have been optically detected. In most cases the companions are white dwarfs. They are however extremely faint ($M_V \simeq 26$). The radius of the white dwarf can be estimated directly from the optical flux and the mass estimated using the theoretical mass-radius relation. Additional information can be gained from the surface gravity by fitting a model atmosphere. White dwarf companions of millisecond pulsars are believed to be He stars with masses $M_2 < 0.5 M_\odot$. Atmospheres of He white dwarfs are however not well understood. Two pulsars have been found in binaries with main sequence companions (B1259–63 with a Be star companion and J0045–7319 with a B1V star companion).

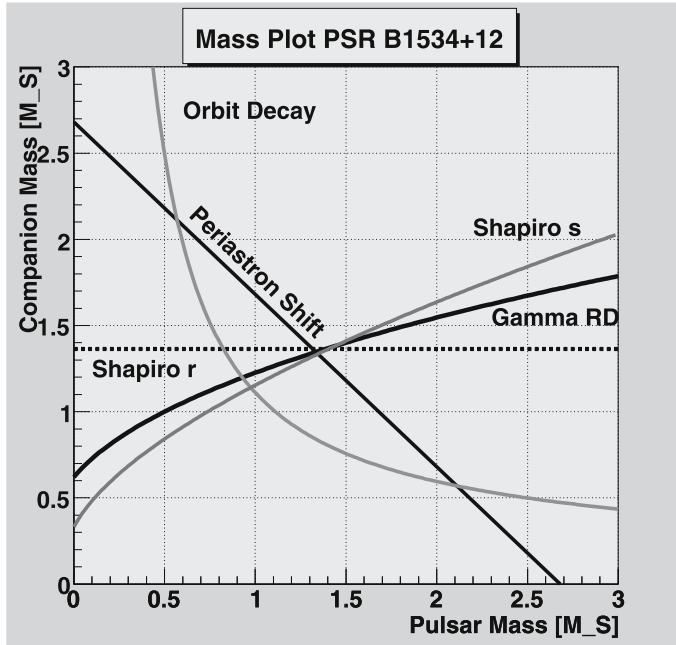


Fig. 6.30. Mass diagram for the binary pulsar system B 1534+12. The five post-Keplerian parameters $\dot{\omega}$, γ_{RD} , r , s and \dot{P}_b give constraints on the two masses. The offset in the orbital decay is due to gravitational acceleration effects in the globular cluster

6.5.5 The Double Pulsar System PSR 0737-3039A+B

The pulsar PSR J0737–3039 was discovered as part of a high-latitude multibeam survey of the southern sky using the 64-m Parkes radio telescope. It was found to be in a 2.4 hr eccentric orbit that the observed orbital parameters suggested was another neutron star [245]. Lyne et al. [261] reported the detection of the 2.8 s pulsar as the companion to a 23 ms pulsar as the first double pulsar system (Table 6.7). Also short eclipses and orbital modulation of the radio fluxes are observed. The relativistic periastron advance, $\dot{\omega}$, and the Doppler factor γ_{RD} could be measured, and also the Shapiro delay in the pulse arrival times of A due to the gravitational field of B could be detected (Fig. 6.32). This provides four measured post-Newtonian parameters resulting in the mass plot Fig. 6.32.

The detection of the pulsar B opens new constraints for this particular system. First, we can exclude all regions in the mass plot due to the requirement $\sin i \leq 1$. Secondly, with a measurement of the semimajor axes of orbit A and B, we obtain a precise measurement of the mass ratio, $R(M_A, M_B) = M_A/M_B = x_B/x_A$, providing a further constraint in the mass plot. The R line is independent of any strong field effects. These constraints make this system the most overdetermined DNS binary to date and a truly unique lab for testing relativistic gravity. The position of the allowed region in Fig. 6.32 also determines the inclination of the orbit with

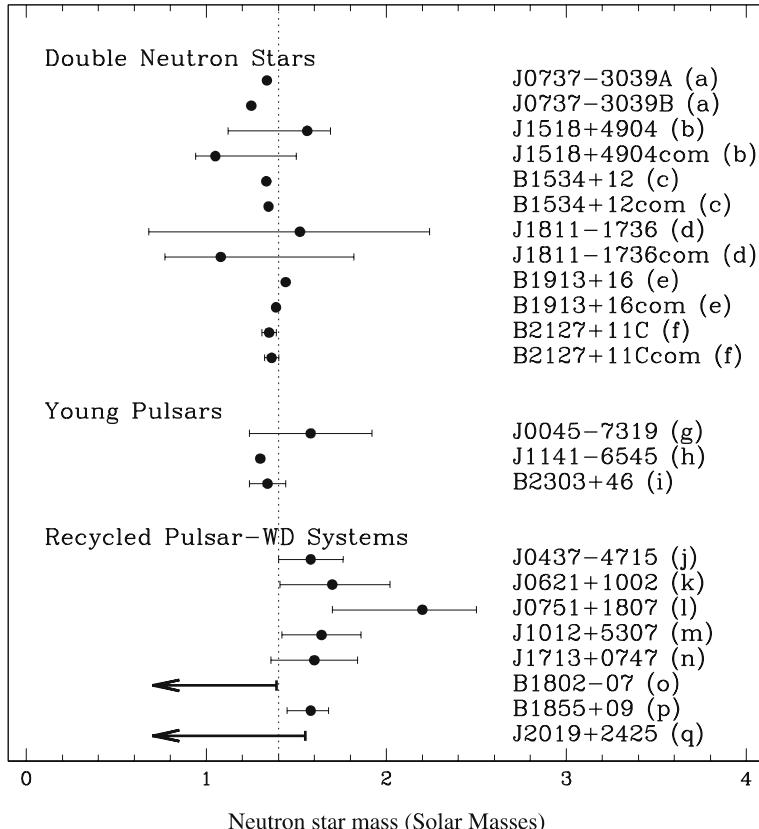


Fig. 6.31. Neutron star masses from observations of radio pulsar systems. Six double neutron star systems are shown at the top of the diagram. Three young pulsars are companions of B stars. Finally, we show recent data from radio pulsars with recycled white dwarf companions. Vertical line is drawn at $M = 1.40 M_{\odot}$

respect to the line-of-sight. This system is observed nearly edge-on with an angle $i \simeq 87$ degrees. This is very favorable for the observation of the Shapiro time delay (Fig. 6.33).

The formation of double neutron star binaries can be understood by a binary evolution scenario which starts with two main sequence stars. The initially more massive star evolves first and explodes in a supernova to form a neutron star. Under favorable conditions, this neutron star remains bound to its companion and spins down as a normal pulsar for the next few million years. At some later time, the secondary star comes to its end of its main-sequence lifetime and begins the red giant phase. Depending on the orbital parameters of the system, an accretion disk is formed around the neutron star with plasma from the red giant, making the system now visible as an X-ray binary. The accretion process transfers angular momentum to the neutron star, spinning it up to short periods and dramatically reducing its

Table 6.7. Timing data for the first double pulsar relativistic binary system PSR 0737–3039A+B (Lyne et al. [261]). The numbers in brackets give the errors. In the first group, you find the astrometric parameters, the second group gives the classical orbital elements of binary systems, and the third group lists the post-Keplerian parameters. Finally, we give some important derived parameters

Pulsar	PSR J0737–3039A	PSR J0737–3039B
Spin period P [ms]	22.69937855615(6)	2773.4607474(4)
Period derivative \dot{P}	$1.74(5) \times 10^{-18}$	$0.88(13) \times 10^{-15}$
Epoch of period (MJD)	52870.0	52870.0
Right ascension (J2000)	07:37:51.247(2)	–
Declination (J2000)	−30:39:40.74(3)	–
Dispersion [pc cm ^{−3}]	48.914(2)	48.7(2)
Timing accuracy [μ s]	27	2660
Orbital period P_b [d]	0.102251563(1)	–
Eccentricity e	0.087779(5)	
Semimajor axis $a_p \sin i/c$ [s]	1.41504(2)	1.513(4)
Periastron length ω_0 [deg]	73.805(3)	73.805 + 180.0
Periastron passage T_0 [MJD]	52870.0120589(6)	–
Periastron shift $\dot{\omega}$ [deg/yr]	16.90(1)	–
Grav/Doppler effect γ_{RD} [ms]	0.38(5)	
Shapiro time delay r [μ s]	5.6(−12,+18)	
Orbital inclination $s = \sin i$	0.99995(−32,+4)	
Orbit decay \dot{P}_b [10 ^{−12} s/s]	not measured	
Characteristic age [Myr]	210	50
Surface magnetic field [gauss]	6.3×10^9	1.6×10^{12}
Spin-down luminosity \dot{E} [erg/s]	5.8×10^{34}	1.6×10^{30}
Mass function [M_\odot]	0.29097(1)	0.356(3)
Distance [kpc]	0.6	
Total mass [M_\odot]	2.588(3)	
Mass ratio m_A/m_B	1.069(6)	
Orbital inclination [deg]	87(3)	
Geodetic precession period [yr]	75	71

magnetic field strength. A limiting spin period is then obtained due to equilibrium between the magnetic pressure of the accreting neutron star and the disk pressure of the infalling matter.

The orbital separation of such systems is dramatically reduced when matter from the secondary is expelled from the system, resulting in a very compact system consisting of a helium star and a neutron star. A sufficiently massive helium star will undergo a supernova explosion forming a young second neutron star. If the stars remain bound, the resulting system is a pair of neutron stars in eccentric orbit with different magnetic field strengths, as is the case in the observed system. With masses $M_A = 1.34 M_\odot$ and $M_B = 1.25 M_\odot$, neutron star A shows a typical mass, while

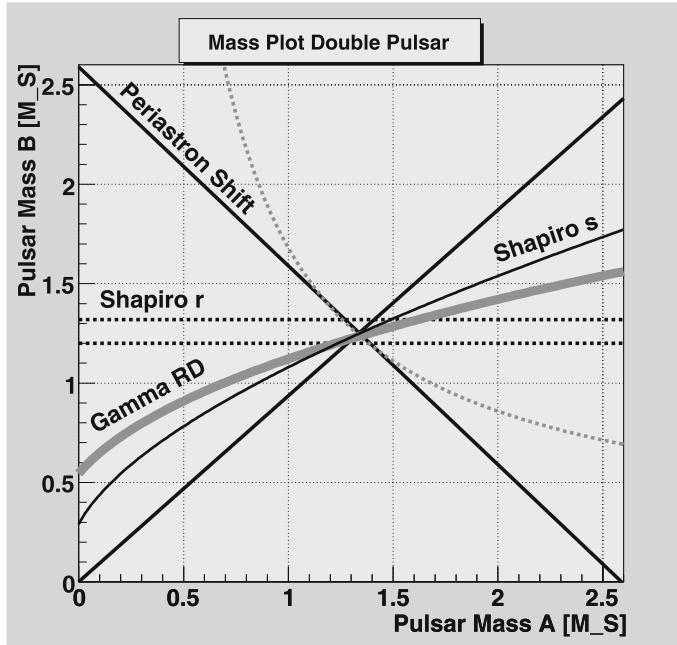


Fig. 6.32. Mass diagram for the binary pulsar system PSR 0737–3039A+B. The mass ratio $R = m_A/m_B = x_B/x_A$ provides the strongest constraint and determines together with the periastron advance uniquely the masses and the inclination. Based on these data, the orbital decay is predicted to be $\dot{P}_b = -1.24 \times 10^{-12}$ (dotted line). Data are taken from [244]

neutron star B has a somewhat smaller mass than the others (Fig. 6.31). The time since the second supernova explosion can be estimated by comparing the spin-down age of the pulsars, $\tau_A \simeq 4\tau_B$.

This system has other interesting properties. The separation of the two pulsars in their orbits is typically 900,000 km, or 3 lightseconds. The large orbital inclination means that, at conjunction, the line-of-sight to one pulsar passes within about 0.15 lt-sec of the other. This is smaller than the light cylinder radius $R_L = c/\Omega = 0.45$ lt-sec of pulsar B, however much greater than the light cylinder of pulsar A. As the pulsar moves, the line-of-sight from A passes through the magnetosphere of pulsar B, providing the opportunity to probe the physical conditions in the magnetosphere of pulsar B (see Fig. 6.41). This allows us to probe the radio transmission properties, the plasma density and the magnetic field structure of the magnetosphere. In fact a short occultation of the pulses from A has already been measured centered upon superior conjunction (Lyne et al. [261]). The duration of the occultation is only 20–30 s. The strong pulsar wind from pulsar A (about a factor one hundred stronger than the wind from B) will form a kind of magnetopause around the magnetosphere of pulsar B (similar to the Earth’s magnetopause formed by the solar wind), and this will be probably located within the light cylinder of B.

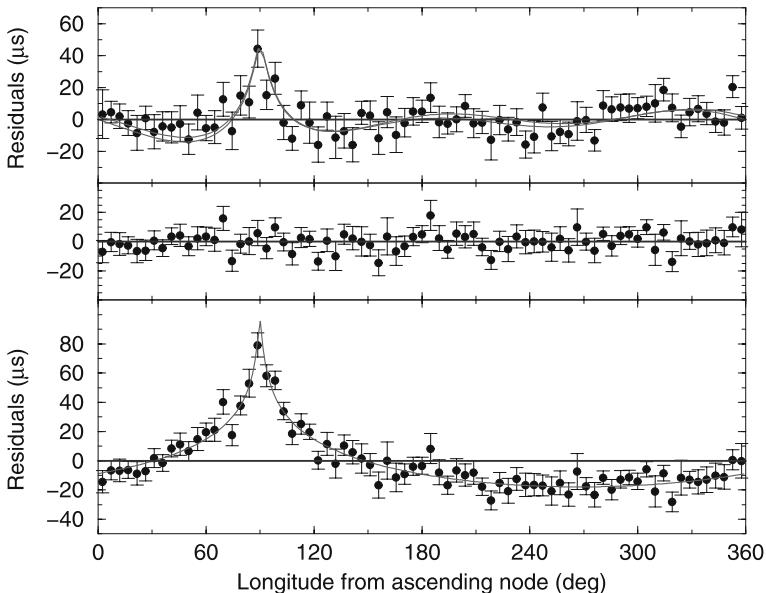


Fig. 6.33. The effect of the Shapiro time delay caused by the gravitational potential of B seen in the timing residuals of the double pulsar. *Top:* timing residuals obtained by subtracting the full timing model; *bottom:* timing residuals obtained by setting the Shapiro delay parameters $r = 0 = s$. Figure adapted from Kramer et al. [245]

The pulsar wind from A is forced to flow around the quenched magnetosphere of B.

6.6 Neutron Stars in our Galaxy

Neutron stars in isolation or as companions of compact objects can appear as radio and/or X-ray pulsars. As of 2005, more than 1500 radio pulsars have been detected in large radio surveys⁴ and about 150 neutron stars are known as X-ray sources.

6.6.1 100 Million Neutron Stars in the Galaxy

Neutron stars are thought to be born in type II supernovae. The typical event rate for such supernovae in spiral galaxies is of the order of one event every 50–100 years, though the historical event rate in our own Galaxy is somewhat short of this number due to various reasons. This means that our Galaxy is populated by at least a few hundred million neutron stars. Since the lifetime of a typical pulsar is only a few million years, we estimate to have about 10,000 active radio pulsars in our Galaxy.

⁴ See the on-line pulsar catalogue: www.atnf.csiro.au/research/pulsars/psrcat.

Of these objects, only some fraction is visible due to the lighthouse effect. This demonstrates that the total number of detectable pulsars will be in the range of a few thousand objects for the entire Galaxy, depending somewhat on the topology of the magnetosphere of neutron stars. Dead pulsars will live forever as cooling rockets when flying through the Galaxy. The neutron star RX J1856–37 recently detected by HST is an example of such a flying graveyard. It is located at a distance of 200 lightyears and moves with a speed of 100 km/s through space. Its surface temperature is about 700,000 K.

Similar to white dwarfs, we can distinguish between three types of neutron stars:

- **Thermal emission from isolated neutron stars:** Observations of isolated neutron stars (or neutron stars in quiescent X-ray binaries) are extremely important in fundamental physics, as thermal emission from the surface of a neutron star carries signatures of its gravitational field, which may be used to infer its mass and radius. Detection of absorption lines corresponding to elements on the neutron star atmosphere and measurement of their gravitational redshift would provide rather accurate data. From the gravitational redshift at the surface of the neutron star, the ratio between its mass and radius may be measured, providing a very strong constraint on neutron star models. Such models give physics an experimentally testable handle on properties of matter at (supra-)nuclear densities.
- **Rotation-powered neutron stars (radio pulsars):** Neutron stars rotate very rapidly, up to 600 times per second. But how are they spinning when they are born? They may be born rotating very fast, with periods comparable to a millisecond (although evidence is ambiguous). After that, they spin down ever after because of magnetic torques. This seems to be supported by the fact that some of the youngest pulsars, such as the Crab pulsar (33 ms) and the Vela pulsar (80 ms) have unusually short periods. After a pulsar is born, its magnetic field will exert a torque and slow it down, with typical spin-down rates of 10^{-13} s/s for a young pulsar like the Crab.

Neutrons were discovered in 1932, and very shortly afterward (in 1934) a suggestion was made by Walter Baade and Fritz Zwicky that neutron stars were formed in supernovae. But for many decades after that, neutron stars were just hypothetical phenomena that did not attract much interest. Since the stars are so small, people felt that the prospects for observing them were minimal, and thus little effort was expended on theory or observation of neutron stars.

This changed dramatically in 1967, due to serendipity and the diligence of an English graduate student by the name of Jocelyn Bell. Bell and her advisor, Anthony Hewish, were working on radio observations of quasars, which had been discovered in 1963. Bell and some other graduate students constructed a scintillation array for the observations, then she got down to examining the charts of data produced (she had to analyze the miles of charts by hand, since this was in the days before powerful computers). One day she noticed a bit of “scruff” that appeared on the charts every second and a third. The scruff was so regular that she first thought it must be artificial. However, careful checking showed that indeed the signal was extraterrestrial, and in fact that it must be from

outside the Solar System. This source, CP 1919, was the first radio pulsar to be discovered.

The discovery initiated a storm of activity that has still not abated. A number of other pulsars were discovered, including one in the Crab Nebula (Fig. 1.8), site of a famous supernova in the year 1054 that was observed by Chinese, Arabic, and North American astronomers (but not recorded, as far as we know, by Europeans). Within a year or so of the initial discovery, it became clear that (i) pulsars are fast, with periods known in 1968 from 0.033 seconds (the Crab pulsar) to about 2 seconds, (ii) the pulsations are very regular, with a typical rate of change of only a second per 10 million years, and (iii) over time, the period of a pulsar always increased slightly.

With this data, it was realized quickly that pulsars had to be rotating neutron stars. With certain exceptions that do not apply in this case, if a source varies over some time t , then its size must be less than the distance light can travel in that time, or ct (otherwise the variation would be happening faster than the speed of light). Thus, these objects had to be less than 300,000 km/s times 0.033 seconds, or 10,000 km, in size. This restricts us to white dwarfs, neutron stars, or black holes. You can get a periodic signal from such objects via pulsation, rotation, or a binary orbit. White dwarfs are large enough that their maximum pulsational, rotational, or orbital frequencies are more than a second, so this is ruled out. Black holes do not have solid surfaces to which to attach a beacon, so rotation or vibration of black holes is eliminated. Black holes or neutron stars in a binary could produce the required range of periods, but the binary would emit gravitational radiation, the stars would get closer together, and the period would decrease, not increase (and would do so very quickly, too!). Pulsations of neutron stars typically have periods of milliseconds, not seconds. The only thing left is rotating neutron stars, and this fits all of the observations admirably.

- **Accretion-powered neutron stars in X-ray binary systems:** Not all neutron stars are destined to lead a life of isolation. Some of them are born in binaries that survive the supernova explosion that created the neutron star, and in dense stellar regions such as globular clusters some neutron stars may be able to capture companions. In either case, mass may be transferred from the companion to the neutron star.

If the companion star has less than the mass of our Sun, the mass transfer occurs via Roche lobe overflow. If part of the companion star's envelope is close enough to the neutron star, the neutron star's gravitational attraction on that part of the envelope is greater than the companion star's attraction, with the result that the gas in the envelope falls onto the neutron star. However, since the neutron star is tiny, astronomically speaking, the gas has too much angular momentum to fall on the star directly and therefore orbits around the star in an accretion disk. Within the disk, magnetic or viscous forces operate to allow the gas in the disk to drift in slowly as it orbits, and to eventually reach the stellar surface. If the magnetic field at the neutron star's surface exceeds about 10^8 G, then before the gas gets to the stellar surface the field can couple strongly to the matter and force it to flow along field lines to the magnetic poles. The friction of the gas with itself

as it spirals in towards the neutron star heats the gas to millions of degrees, and causes it to emit X-rays.

If the star has a companion, it can accrete from the companion and have its rotational frequency altered that way. If the companion is a low-mass star, say half the mass of our Sun or lower, accretion tends to proceed by Roche lobe overflow. The radius of the inner edge of the disk is determined by the strength of the magnetic field; the stronger the field, the farther out it can control the accretion flow (for a given accretion rate). The star then (more or less) tries to come to equilibrium with the Keplerian angular velocity of the matter at the inner edge of the accretion disk. This means that neutron stars with relatively small (10^8 to 10^9 gauss) magnetic fields can be spun up to high frequencies, and this is the accepted picture of how we get millisecond pulsars.

If the companion of the neutron star is a high-mass star (over 10 solar masses) instead, then the matter that makes it onto the neutron star goes in the form of a low angular momentum wind. Therefore, the neutron star is not spun up to such high frequencies; in fact, some X-ray pulsars that are in high-mass systems have periods longer than 1000 seconds. The process of wind accretion is a very complicated one, and numerical simulations of the process push the limits of computers. It appears that, in some circumstances, a disk may form briefly around the neutron star, only to be dissipated and replaced by a disk going the other way. One barrier to understanding this kind of accretion is that, even with today's computers, high-resolution 3D simulations just aren't feasible now, so we have to derive what insight we can from good two-dimensional calculations.

6.6.2 Thermal Emission from Isolated Neutron Stars

Neutron stars are numerous, but observationally very diverse. In particular, direct emission from the neutron star surface layers (the only source of information on the physical conditions of the star) is only accessible in a handful of sources. This is because the surface of old neutron stars (age $> 10^6$ yr) is too cold to emit X-rays, while the radiation from young active radio pulsars (age $< 10^4$ yr) is dominated by nonthermal emission from the magnetosphere surrounding the star. In the last decade ROSAT satellite observations led to the discovery of seven dim sources, which were later associated with radio-silent, isolated, middle-age (age 10^5 – 10^6 yr) neutron stars. They are characterized by clean thermal emission at energies of about 0.1 keV without any trace of contamination from a surrounding supernova remnant or magnetospheric activity. They are located within a few hundred pc, close enough to be studied with the latest X-ray orbiting observatories and therefore represent important targets for the study of neutron star surface emission.

Data on Cooling Neutron Stars

Observations of neutron stars with detection of thermal emission are summarized in Table 6.8 and 6.9. These tables display four inferred quantities: the spin-down age t_{sd} ,

Table 6.8. Neutron stars with hydrogen atmospheres. Data from [321]

Star	$\log t_{\text{sd}}$ yr	$\log T_{\infty}$ K	Distance kpc	$\log L_{\infty}$ erg/s
RX J0822–4247	3.90	6.24 ± 0.04	1.9–2.5	33.85–34.0
1E 1207.4–5209	5.53	6.21 ± 0.07	1.3–1.9	33.27–33.74
RX J0002+6246	–	6.03 ± 0.03	2.5–3.5	33.08–33.33
PSR 0833–45 (Vela)	4.05	5.83 ± 0.02	0.22–0.28	32.41–32.70
PSR 1706–44	4.24	5.8 ± 0.13	1.4–2.3	31.81–32.93
PSR 0538+2817	4.47	6.05 ± 0.10	1.2	32.6–33.6

Table 6.9. Neutron stars with black-body atmospheres. The uncertainties in the observed radii are still too big for confrontation with theory. Data from [321]

Star	$\log t_{\text{sd}}$ yr	$\log T_{\infty}$ K	R_{∞} km	Distance kpc	$\log L_{\infty}$ erg/s
RX J0822–4247	3.90	6.65 ± 0.05	1–1.6	1.9–2.5	33.60–33.90
1E 1207.4–5209	5.53	6.48 ± 0.01	1.0–3.7	1.3–3.9	32.70–33.88
RX J0002+6246	–	6.15 ± 0.11	2.1–5.3	2.5–3.5	32.18–32.81
PSR 0833–45 (Vela)	4.05	6.18 ± 0.02	1.7–2.5	0.22–0.28	32.04–32.32
PSR 1706–44	4.24	6.22 ± 0.04	1.9–5.8	1.8–3.2	32.48–33.08
PSR 0656+14	5.04	5.71 ± 0.04	7.0–8.5	0.26–0.32	32.18–32.79
PSR 0633+1748 (Gem)	5.53	5.75 ± 0.04	2.7–8.7	0.123–0.216	30.85–31.51
PSR 1055–52	5.43	5.92 ± 0.02	6.5–19.5	0.5–1.5	32.07–33.19
RX J1856.5–3754	–	5.6–5.9	> 16	0.105–0.129	31.44–31.68
RX J0720.4–3125	6.0 ± 0.2	5.55–5.95	5.0–15.0	0.1–0.3	31.3–32.5

the surface temperature T_{∞} , the total thermal luminosity L_{∞} and the estimated distance d . The subscript ∞ refers to quantities observed at Earth which are redshifted relative to their values at the stellar surface. The data are taken from references given in [321]. Table 6.8 refers to models incorporating atmospheres dominated by hydrogen, while Table 6.9 presents properties inferred from black-body atmospheres. There are also sources available with only upper limits.

The estimation of T_{∞} and L_{∞} from the observed spectral fluxes requires atmospheric modelling with the inclusion of three additional factors: the composition of the atmosphere, the column density of X-ray absorbing material between the star and Earth, and the surface gravitational redshift. The column density is important, since the bulk of the emitted flux from neutron stars is absorbed by interstellar hydrogen, before it reaches the Earth (Fig. 6.34). Since no narrow spectral lines are observed in the neutron stars in Table 6.8 and 6.9, the atmospheric composition of these stars is unknown. Neutron star atmospheres can be either light-element (i.e. H and He) or heavy-element dominated. The latter ones have spectral distributions more closely resembling a black-body distribution, due to higher opacities of heavy elements. If the distance is known, the neutron star emission radius can be inferred. It seems that

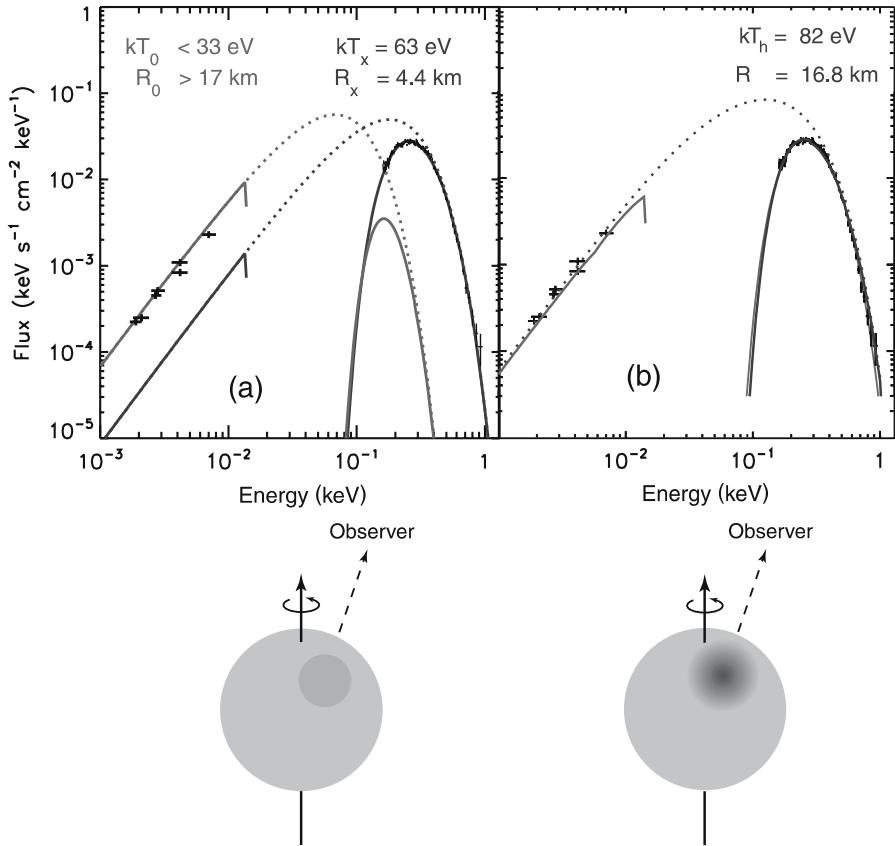


Fig. 6.34. The X-ray spectrum of the isolated neutron star RJ 1856–3754 for (a) a two-component model and (b) a model with continuous temperature distribution. The X-ray and optical data show no spectral features and are best fitted with a two-component black-body model with $kT_{bb,X}^{\infty} \simeq 63.5 \text{ eV}$ and $R_{bb,X}^{\infty} \simeq 4.4(d/120 \text{ pc}) \text{ km}$ for the hot X-ray emitting region, and $kT_{bb,opt}^{\infty} < 33 \text{ eV}$ and $R_{bb,opt}^{\infty} > 17(d/120 \text{ pc}) \text{ km}$ for the rest of the neutron star surface responsible for the optical flux. The optical emission is then the Rayleigh–Jeans spectrum emitted by the neutron star surface and the hot spot. The neutron star radius would be 14 km. This would rule out quark stars. Figure adapted from Burwitz et al. [93] and Trümper et al. [400]

the inferred neutron star radii for stars younger than 100,000 years are consistent with canonical values only if they are modelled with light-element atmospheres. On the other hand, stars older than 100,000 years have inferred radii close to the canonical range only if modelled with heavy-element atmospheres. It seems therefore that the atmospheric composition of a neutron star evolves from light elements to heavy elements on a time-scale of about 100,000 years.

Ages for the observed cooling neutron stars can be estimated from the observed spin-down rate $t_{\text{sd}} = P/2\dot{P}$, where P and \dot{P} are the period and its time deriva-

tive, respectively. In some cases, one also has kinematic information by relating pulsar transverse velocities to distances from the presumed sites of origin (e.g. OB associations). The distances are estimated from pulsar dispersion measures, estimated distances for the corresponding supernova remnant, or by observations of the interstellar absorption to nearby stars.

These dim isolated neutron stars (DINs) are key objects in compact object astrophysics. They offer a unique laboratory for investigating the properties of matter under extreme conditions, such as the equation of state at supranuclear densities, or the interaction of highly relativistic plasmas with radiation in the presence of giga- or teragauss magnetic fields. Detailed X-ray spectra of DINNs have been recently obtained with Chandra and XMM–Newton, and show quite unexpected characteristics. The prototype of the class, RX J1856.5–3754, exhibits a featureless spectrum extremely close to a pure black-body (Fig. 6.34). Broad absorption features have been detected in four pulsating sources with evidence of a spectral variation with phase. Very recently, spectral evolution on timescale of years have been reported for the second most luminous source, RX J0720.4–3125. In addition, when detected the optical counterpart lies a factor about 5–10 above the extrapolation of the X-ray black-body at optical wavelengths. All these new findings represent a challenge for conventional atmospheric models, typically based on surface temperature distributions induced by a dipolar magnetic field.

Compact Stellar Radii from Thermal Emission

The observed temperature $T_{\text{eff},\infty}$ determines the observed luminosity

$$L_\infty = 4\pi\sigma_{SB} T_{\text{eff},\infty}^4 R_\infty^2 = L_* \alpha^2(R). \quad (6.293)$$

The last equality follows from the fact that the energy is redshifted and time is also redshifted with respect to the timescale at the surface of the star, $dt_\infty = d\tau/\alpha(R)$. The BB spectrum is redshifted when it leaves the gravitational field of the star, i.e.

$$T_{\text{eff},\infty} = \alpha(R) T_{\text{eff}} = \sqrt{1 - 2GM/c^2R} T_{\text{eff}}. \quad (6.294)$$

Since $L_* = 4\pi R^2 \sigma_{SB} T_{\text{eff}}^4$, this has the consequence that the *observed radius* R_∞ is *apparently bigger than the true radius of the neutron star*

$$R_\infty = \frac{R}{\alpha(R)} \simeq \frac{R}{\sqrt{1 - 2GM/c^2R}}. \quad (6.295)$$

An intrinsic radius for a neutron star of 1.4 solar masses corresponds to an observed radius between 13 and 14.5 km (Fig. 6.35). If massive neutron stars would have quark cores, we would expect observed radii between 12 and 13 km.

Neutron Star Cooling

Neutron stars are formed at much higher temperatures than white dwarfs, typically at 50 MeV. Much of this initial thermal energy is radiated away from the interior

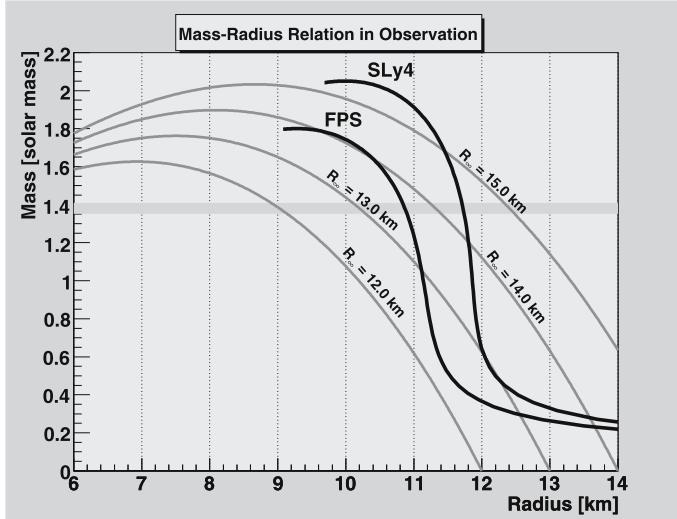


Fig. 6.35. Radius constraints on the mass–radius relation are given by *lines* of constant observed radii R_∞ , together with the mass–radius relation for various EoS

of the star by means of neutrino emission (mainly by the modified URCA process and neutrino Bremsstrahlung). After one day, the neutron star has cooled down to a few billion K, and after 100 years the star’s interior becomes nearly isothermal (thermally relaxed). After this period, the energy balance of the cooling neutron star is determined by

$$\frac{dE_{th}}{dt} = C_V(T_i) \frac{dT_i}{dt} = -L_\nu(T_i) - L_\gamma(T_*) + L_H, \quad (6.296)$$

with the following luminosities (in Newtonian expressions)

$$C_V = \frac{4\pi}{3} R^3 c_v T \quad (6.297)$$

$$L_\nu = \frac{4\pi}{3} R^3 Q_\nu T^8 \quad (6.298)$$

$$L_\gamma = 4\pi R^2 \sigma_{SB} T_*^4 \propto T^{2+a} , \quad a \ll 1. \quad (6.299)$$

This Newtonian equation will be modified by some relativistic corrections (see below). T_i is the internal (core) temperature, T_* the surface temperature, and L_H denotes heating processes (e.g. frictional heating by superfluid neutrons in the inner crust), $L_*(T_*) = 4\pi R_*^2 \sigma T_*^4$, and $L_\nu(T_i) = \int \epsilon_\nu dV$ is the neutrino luminosity. For a recent review on the details, see Yakovlev and Pethick [424, 425].

The above Newtonian equations have to be replaced by the correct relativistic relations for spherical stars (see problem 6.10)

$$\frac{\exp(-\lambda - 2\Phi)}{4\pi r^2} \frac{\partial}{\partial r} (\exp(2\Phi)L_r) = -Q_v + Q_h - \frac{c_v}{\exp(\Phi)} \frac{\partial T}{\partial t} \quad (6.300)$$

$$\frac{L_r}{4\pi\kappa r^2} = \exp(-\lambda - \Phi) \frac{\partial}{\partial r} (\exp(\Phi)T) . \quad (6.301)$$

Q_v is the neutrino emissivity, c_v is the heat capacity per unit volume, κ is the thermal conductivity, and L_r is the local luminosity given by the nonneutrino heat flux transported through a sphere of radius r . Q_H denotes the energy production by reheating (e.g. dissipation of rotational energy). After thermal relaxation, the redshifted temperature $T_i(t) = \exp(\Phi)T(t, r)$ becomes constant throughout the stellar interior. Then we obtain the equations of global thermal balance (6.296) with the interpretation

$$L_v(T_i) = \int dV Q_v(T) \exp(2\Phi) \quad (6.302)$$

$$C_V(T_i) = \int dV c_v(T) , \quad (6.303)$$

where $dV = 4\pi r^2 / \sqrt{1 - 2Gm(r)/c^2r}$ is the element of proper volume.

Neutrino Emission Processes

Neutrino emission is generated by various reactions in the interior of neutron stars. Neutrinos escape and carry away energy. The neutrino mechanisms in the core can be subdivided into *slow* and *fast* processes with emissivities given by

$$Q_{\text{slow}} = Q_s T_9^8, \quad Q_{\text{fast}} = Q_f T_9^6 , \quad (6.304)$$

where Q_s and Q_f are slowly varying functions of the density, presented in Table 6.10. The most powerful neutrino emission is provided by the *direct Urca* processes in nucleon and hyperon matter. It consists of a pair of reactions, the beta decay of a neutron and electron capture on a proton, whose net effect is the emission of a neutrino–antineutrino pair. The composition of the matter remains unchanged. This process can only occur if the proton concentration is sufficiently high. In degenerate matter, only particles with energies within $\simeq k_B T$ of the Fermi surface can participate in reactions, since other processes are blocked by the Pauli exclusion principle. If the proton and electron Fermi momenta are too small compared with the neutron Fermi momenta, the process is forbidden, because momentum conservation cannot be satisfied. One finds that the ratio of the number density of protons to that of nucleons must exceed 0.1 for this process to be allowed. Proton fractions in the outer core are estimated to be lower than this. It could only occur in inner hyperon cores. In particular, a concentration of Λ hyperons of order 10^{-3} could lead to rapid neutrino emission.

Pion-condensed and kaon-condensed matter could also lead to fast cooling processes. Direct Urca processes can operate on the nucleon quasiparticles, which, in the case of a pion condensate, is a coherent superposition of a neutron and a proton.

Table 6.10. Neutrino emission processes in the cores of neutron stars and their emissivities. T_9 is the core temperature in units of 10^9 K. α_c denotes the strong coupling constant for quarks

Process Name	Process	Emissivity Q_ν (erg cm $^{-3}$ s $^{-1}$)	Reference
Modified Urca	$n + n' \rightarrow n' + p + e^- + \bar{\nu}_e$	$\simeq 10^{20} T_9^8$	Friman and Maxwell [162]
	$n + p + e^- \rightarrow n' + n + \nu_e$		
Direct Urca	$n \rightarrow p + e^- + \bar{\nu}_e$	$\simeq 10^{27} T_9^6$	Lattimer et al. [248]
	$p + e^- \rightarrow n + \nu_e$		
Quark Urca	$d \rightarrow u + e^- + \bar{\nu}_e$	$\simeq 10^{26} \alpha_c T_9^6$	Iwamoto [215]
	$u + e^- \rightarrow d + \bar{\nu}_e$		
Kaon condensate	$n + K^- \rightarrow n + e^- + \bar{\nu}_e$	$\simeq 10^{24} T_9^6$	Brown et al. [91]
	$n + e^- \rightarrow n + K^- + \nu_e$		
Pion condensate	$n + \pi^- \rightarrow n + e^- + \bar{\nu}_e$	$\simeq 10^{26} T_9^6$	Maxwell et al. [276]
	$n + e^- \rightarrow n + n' + \nu_e$		

Slow neutrino reactions operate everywhere in the core, in particular in the outer core (hence especially in low-mass stars). For matter consisting only of neutrons, protons and electrons, these are the modified URCA process and NN -Bremsstrahlung. The modified Urca process differs from the direct one by an additional spectator nucleon N required to ensure conservation of momentum and energy. There are three Bremsstrahlung processes in npe matter (nn , np and pp). With hyperons or quarks, other modified Urca processes could occur.

The neutrino reactions are drastically affected by baryon superfluidity. When the temperature drops much below the critical temperature T_c of a given species, the energy gap in the baryon spectrum makes these baryons inactive by suppressing all reactions involving such baryons. As an example, superfluid protons in npe matter suppresses all Urca processes, but does not affect neutron–neutron Bremsstrahlung.

Heat Capacity of Neutron Stars

The major contribution to the heat capacity in a neutron star comes from the core. It is the sum of the heat capacities of the various degenerate constituents of dense matter. The heat capacity per unit volume of a normal (nonsuperfluid) particle is given by

$$c_v = N(0) \frac{\pi^2}{3} k_B^2 T = \frac{m^* n}{p_F^2} \pi^2 k_B^2 T, \quad (6.305)$$

where p_F is the Fermi momentum, $N(0)$ the density of states at the Fermi surface, and m^* the effective mass at the Fermi surface (see Landau–Lifschitz). These quantities are given through the relations

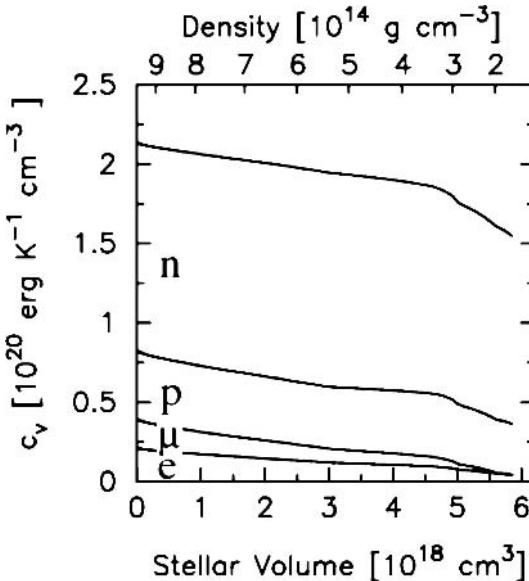


Fig. 6.36. The contributions of the various components for the specific heat in a minimal model of neutron star cores (not including any hyperons) [321]. The crust–core transition occurs at about half nuclear density

$$N(0) = \frac{3n}{p_F v_F}, \quad v_F = \frac{\partial \epsilon}{\partial p} \Big|_{p_F}, \quad (6.306)$$

where v_F denotes the velocity at the Fermi surface.

Neutron star cooling is strongly affected by superfluidity of nucleons in stellar interiors. In principle, the EoS and the superfluidity model have to be obtained from the same nuclear Hamiltonian. Such calculations are however extremely difficult. Pairing interactions have a great influence on the specific heat. When T reaches T_c , there is a sharp increase in the specific heat due to the large fluctuations occurring in a second-order phase transition. Subsequently, when $T \ll T_c$, a Boltzmann suppression occurs due to the presence of a gap in the energy spectrum (for a discussion of this point, see [321]). Microscopic theories predict two main types of superfluidity inside neutron star cores: a singlet-state proton pairing and a triplet-state neutron pairing at higher densities. These theories still give a large scatter of critical temperatures depending on the nucleon–nucleon interaction model. Note that superfluidity induces an additional neutrino emission process associated with Cooper pairing of nucleons

The main contributions to the npe core comes from neutrons, while the contributions from electrons and protons is only a small contribution (Fig. 6.36). The total thermal energy of a nonsuperfluid neutron star (without quarks) is estimated as $E_{\text{th}} \simeq 10^{48} T_9^2$ erg. The heat capacity of an npe neutron star core with strongly superfluid neutrons and protons is determined by the electrons, which are not superfluid, and is about 20 times lower than for a neutron star with a nonsuperfluid core.

Neutrino vs. Photon Cooling

The basic features of the thermal evolution of a neutron star follow from the equation (6.296). In the neutrino cooling era, $L_\nu \gg L_\gamma$, we find then

$$\frac{dT}{dt} = \frac{q_\nu}{c_v} T^7 \quad (6.307)$$

which has the solution

$$t - t_0 = A \left(\frac{1}{T^6} - \frac{1}{T_0^6} \right), \quad (6.308)$$

where T_0 is the initial temperature at t_0 . This gives for $T \ll T_0$ a behavior $T \propto t^{-1/6}$, or a surface temperature $T_* \propto t^{-1/12}$ (see below). The very slow decay of the temperature is a direct consequence from the high exponent in the neutrino emissivity.

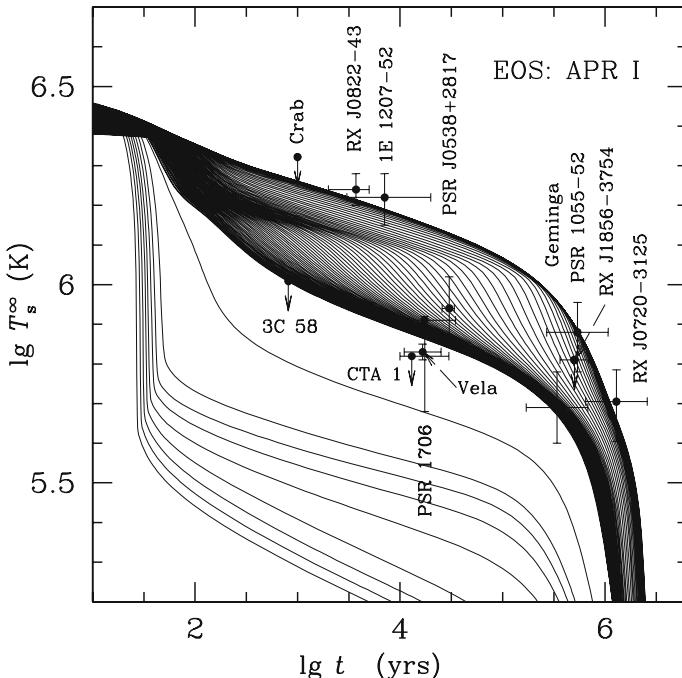


Fig. 6.37. Cooling curves calculated with the APR EoS, including proton pairing in the outer crust and neutron pairing in the inner crust. The various *curves* refer to neutron star masses from 1.01 to $1.92 M_\odot$ (step-size is $0.01 M_\odot$). Included are observed temperatures, and upper limits for 12 isolated neutron stars. The sharp boundary between densely and rarely covered domains corresponds to the direct Urca threshold for massive neutron stars. Figure adapted from [185]

The photon luminosity can be written as

$$L_* = 4\pi R^2 \sigma_{SB} T_*^4 = S T^{2+4a}, \quad (6.309)$$

where the effective temperature T_* has been converted into the internal temperature T through an envelope model with a power-law dependence, $T_* \propto T^{1/2+a}$ (see Fig. 6.39). In the photon cooling era, $L_\gamma \gg L_\nu$,

$$t - t_1 = \frac{B}{4aS} \left(\frac{1}{T^{4a}} - \frac{1}{T_1^{4a}} \right), \quad (6.310)$$

where T_1 is the temperature at t_1 . This last equation shows that the temperature as a function of time is a very steep function in the photon cooling era, $T \propto t^{-1/4a}$ and therefore $L_* \propto t^{-2/a}$, while $T \propto t^{-1/6}$ in the neutrino cooling era, i.e. $T_* \propto t^{-1/8a}$, and therefore $L_* \propto t^{-1/3}$ (see Fig. 6.38).

Since $a \ll 1$, we see that, during the photon cooling era, the evolution is very sensitive to the nature of the envelope and to changes in the specific heat, as induced, e.g. by nucleon pairing. The transition from neutrino dominated cooling towards photon dominated cooling occurs at an age of about 100,000 years.

The Crust–Core Relation

For an outer crust composed of iron, a simple relation between surface temperature and core temperature can be found (Fig. 6.39)

$$T_* = 3.1 \times 10^6 \text{ K} \left(\frac{g_*}{10^{14} \text{ cm s}^{-2}} \right)^{1/4} \left(\frac{T_i}{10^9 \text{ K}} \right)^{0.549}. \quad (6.311)$$

g_* is the gravitational acceleration at the stellar surface. Neutrino emission is the dominant cooling mechanism for the first 100,000 years (Fig. 6.37). After this time, photon emission from the surface of the star gains over neutrino losses, and the surface temperature drops rapidly. In the standard cooling scenario, the modified URCA process $n + N \rightarrow p + N + e + \bar{\nu}_e$ brings the temperature gradually down to about one million K after the neutrino phase. Accelerated cooling occurs in this phase for pion condensations or quark–gluon cores.

Thus, the thermal evolution of neutron stars between 10 years and 10 million years is very sensitive to the composition of the interior. Since the typical temperatures are expected to be in the range of UV–soft-X-ray regime, a hunt for such objects has already been started with ROSAT. In fact thermal radiation has been observed from about a dozen of isolated neutron stars (Fig. 6.38). These data must be, however, considered as upper limits, since heating due to magnetospheric processes could be important. In addition, the strong magnetic fields could lead to a nonuniform surface temperature distribution. As with white dwarfs, the atmosphere of the neutron star also plays an important role for the emergent spectrum. Particularly interesting are the spectra emitted by iron atmospheres which show numerous spectral features produced by iron ions in various ionization stages.

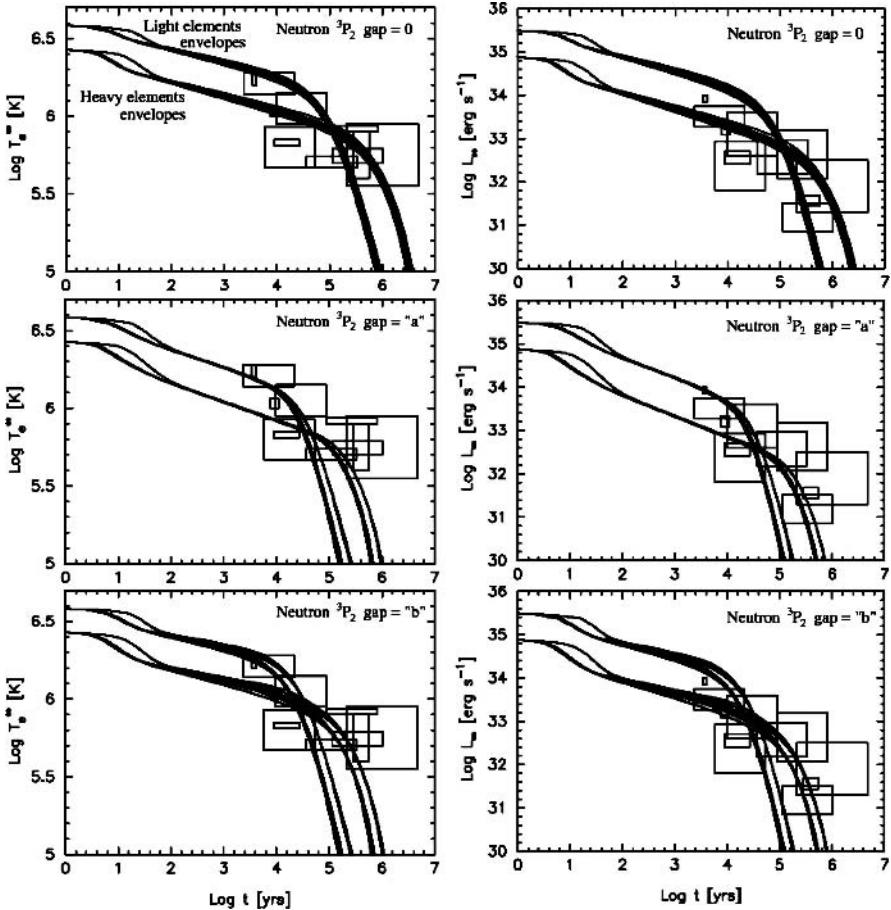


Fig. 6.38. Cooling curves in a minimal model of neutron stars for various compositions of the envelope, all models being $1.4 M_{\odot}$ stars built with the EoS APR98 [321]. The two sets of curves correspond to the extreme models of envelope chemical composition: light elements or heavy elements. The different panels correspond to different assumptions about gap energies. Figure adapted from [321]

Cooling of Neutron Stars with Color Superconducting Quark Cores

The formation of a condensate (K^- or π^-) in the first cooling phase would lead to a rapid increase in the cooling. This is certainly in conflict with present cooling observations. Superconducting quark matter will also affect the cooling curves (Fig. 6.40).

Within nonlocal chiral quark models [81], the critical density for a phase transition can be low enough to occur in stars with masses beyond 1.3 solar masses. The detailed structure of the phase diagram in these models still depends on the strength parameter for the diquark coupling. For the 2SC phase stable hybrid star configu-

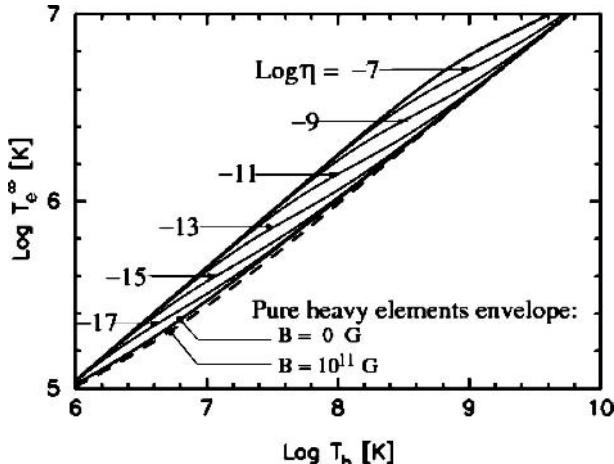


Fig. 6.39. The surface temperature as a function of the core temperature T_b at the bottom of the envelope for various amounts of light elements parametrized by $\eta = \Delta M_L/M$. ΔM_L is the mass in light elements in the envelope. Figure adapted from [321]

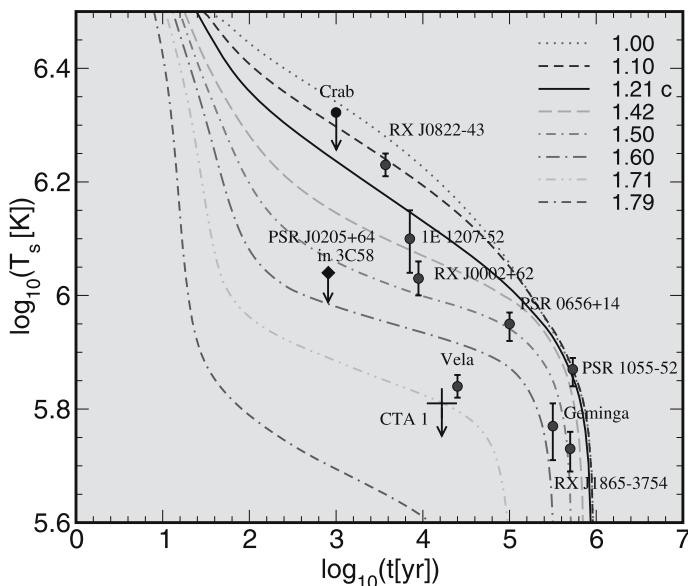


Fig. 6.40. Cooling curves for hybrid neutron stars including a 2SCX quark core. The various curves show the effects for different masses (*labels* in terms of solar masses). Figure adapted from [81]

rations have been found even for masses of 1.3 solar masses. This phase has one unpaired color of quarks (say blue) for which the very effective direct URCA process would occur. This would lead to a too fast cooling of hybrid stars in disagreement with observations. Therefore, it is assumed that a weak pairing channel is operating also for blue colors with a resulting gap of the form

$$\Delta_X(\mu) = \Delta_c \exp(-\alpha(\mu - \mu_c)/\mu_c), \quad (6.312)$$

with μ being the quark chemical potential, $\mu_c = 330$ MeV, $\alpha = 10$ and $\Delta_c = 1.0$ MeV [81]. The physical origin of this X-gap was still unclear, but the inclusion of this gap provides acceptable fits to the cooling curves (Fig. 6.40).

6.6.3 Rotation-Powered Pulsars

The radio pulsars detected so far can be interpreted as rapidly spinning, strongly magnetized neutron stars radiating at the expense of their rotational energy. The energy source for these objects is neither fusion energy, nor gravitational energy. It is simply the rotational energy stored in the rapid rotation of these compact objects.

By measuring the period P and its derivative \dot{P} we can infer some fundamental quantities for these objects. The pulsar slowing down can be written as

$$\boxed{\dot{\Omega} = -k\Omega^n} \quad (6.313)$$

with $n \simeq 3$ as the braking index of the pulsar. This equation can be solved

$$t = -\frac{\Omega}{(n-1)\dot{\Omega}} \left[1 - \frac{\Omega^{n-1}}{\Omega_0^{n-1}} \right]. \quad (6.314)$$

Ω_0 is the initial rotational frequency of the pulsar after its birth in a supernova explosion. For $\Omega \ll \Omega_0$, we get the age of the pulsar (the so-called **spin-down age**)

$$\boxed{t_{\text{sd}} = -\frac{1}{n-1} \frac{\Omega}{\dot{\Omega}} = \frac{1}{n-1} \frac{P}{\dot{P}} = \frac{1}{2} \frac{P}{\ddot{P}}} \quad (6.315)$$

for $n = 3$. The braking index can indeed be measured from the second derivative of the period (presently only available for a few radio pulsars)

$$n = \frac{\Omega \ddot{\Omega}}{\dot{\Omega}^2} = 2 - \frac{P \ddot{P}}{\dot{P}^2}. \quad (6.316)$$

A plot of the observed periods vs. period derivatives is shown in Fig. 6.42, using the ATNF pulsar catalogue⁵. The diagram essentially decays into three subclasses. The bulk of objects is formed by standard pulsars with a mean period of about 0.7 seconds (see Fig. 6.43 for the period histogram) and a typical braking

⁵ See online Web interface: <http://www.atnf.csiro.au/research/pulsars/psrcat>.

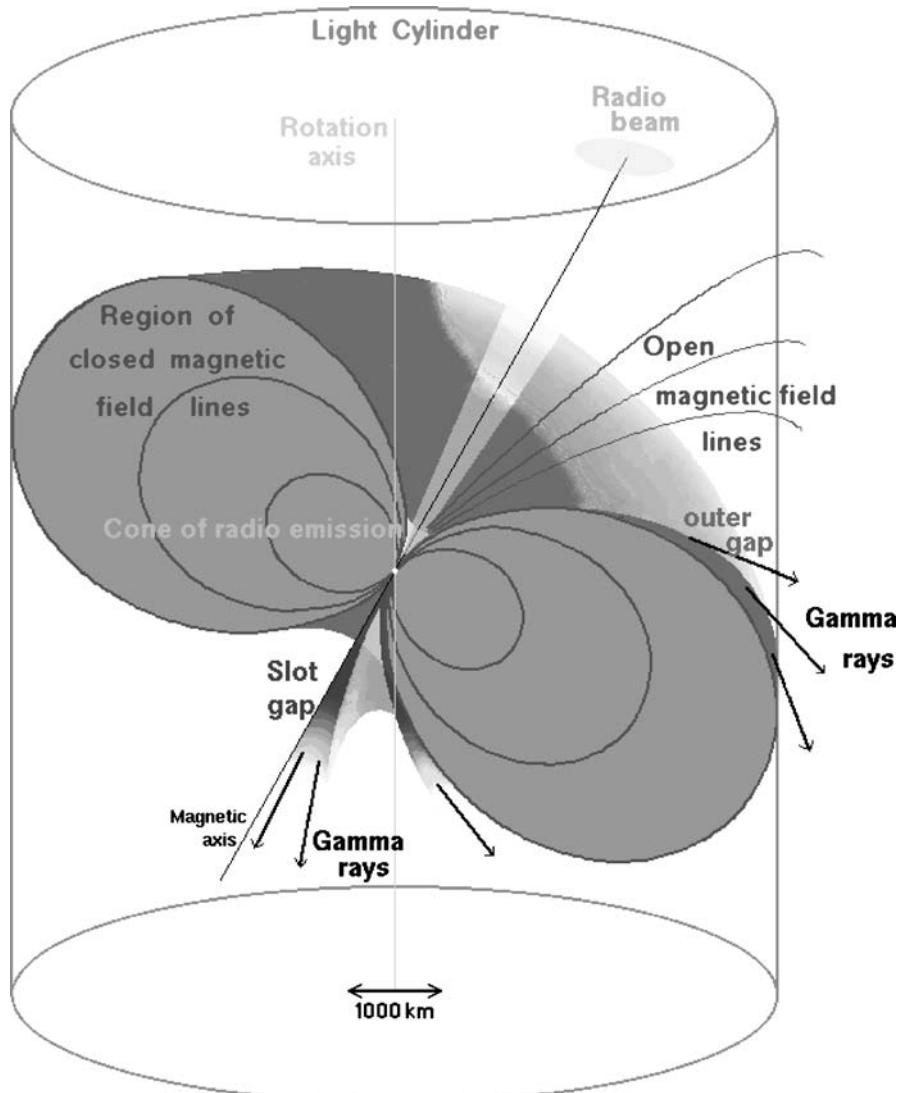


Fig. 6.41. The magnetosphere of rotating neutron stars. **A:** Rotating magnetic field lines define a light cylinder with radius $R_L = c/\Omega \simeq 50,000$ (P/sec) km, when the period P is in units of seconds. At the light cylinder, the dipolar magnetosphere gets distorted. Field lines from the polar caps cross the light cylinder (*open field lines*), plasma generated near the polar cap can escape along these field lines **(B)**. **C:** The polar cap itself acts as a kind of discharge tube, where strong parallel electric fields, generated by the rapid rotation, extract charges from the surface of the neutron star and accelerate them to high energies

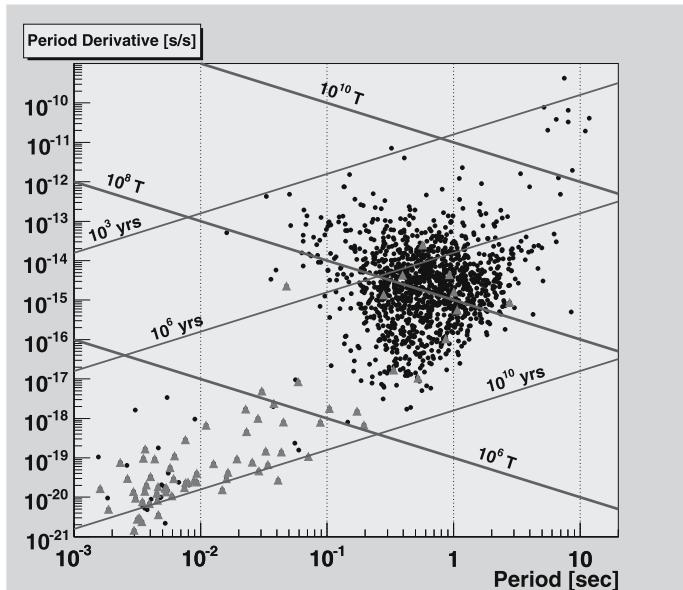


Fig. 6.42. The fundamental pulsar diagram for the ATNF pulsar catalogue. *Triangles* represent pulsars in a binary system. Shown are also lines of constant age and lines of constant surface field strength in units of tesla for the dipole interpretation of the energy loss

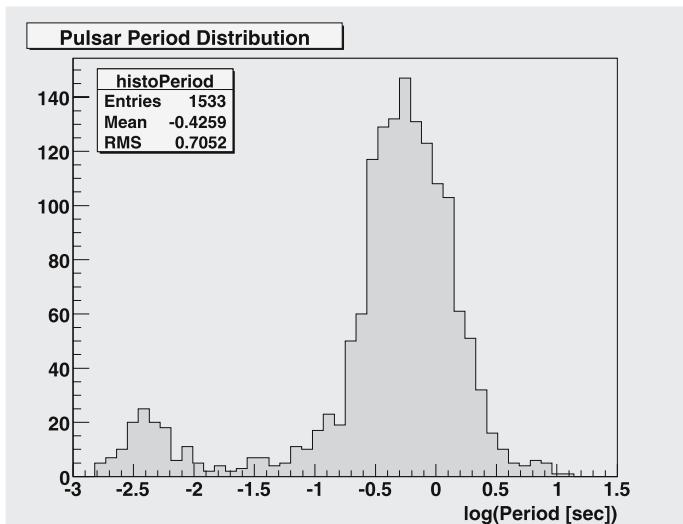


Fig. 6.43. The pulsar period histogram from the ATNF pulsar catalogue

of $\dot{P} = 2.5 \times 10^{-15}$. In the lower left corner we find the millisecond pulsars with periods exceeding 1.5 milliseconds. The objects in the upper right corner represent the soft gamma-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs), which were detected only in recent years. No pulsars have been found in the lower right corner of the diagram, though there are certainly magnetized neutron stars rotating with low frequency. This proves that the pulsar mechanism only operates for certain conditions on period and magnetic field strength.

A rotating vacuum dipole magnetosphere would lose energy according to the dipole formula (the Larmor formula in cgs units)

$$-\dot{E} = \frac{2}{3c^3} \dot{\mu}^2 = \frac{32\pi^4}{3c^3} P^{-4} \mu_{\perp}^2, \quad (6.317)$$

where $\mu_{\perp} = B_* R_*^3 \sin \chi$ is the magnetic moment of the neutron star with surface field strength B_* and radius R_* . The angle χ measures the inclination of the dipole axis with respect to the rotational axis. For this process, the braking index would be three, $n = 3$. Measured values for the braking index are usually below 3, indicating that more complicated plasma processes are involved in the pulsar phenomenon.

Using this expression for the magnetic braking, we can get a handle on the field strength of the neutron star. The rotational loss of the neutron star is given by

$$\dot{E}_{\text{rot}} = I_* \Omega \dot{\Omega} = -4\pi^2 I_* \frac{\dot{P}}{P^3}, \quad (6.318)$$

where I_* is the moment of inertia of the neutron star (Fig. 6.11). Equating the two energy-loss formulae, we obtain an expression for the projected field strength of the neutron star

$$B_* \sin \chi \simeq \sqrt{\frac{3c^3 I_* P \dot{P}}{8\pi^2 R_*^6}} \simeq 10^8 \text{ tesla} \sqrt{P \dot{P}_{-15}}. \quad (6.319)$$

The period P is given in units of seconds and the period derivative \dot{P} in units of $10^{-15} \text{ s s}^{-1}$. The lines of constant field strength are also indicated in the pulsar diagram in Fig. 6.42. According to this interpretation, AXPs and SGRs would have extremely strong magnetospheres with surface field strengths $B_* > 10^{14}$ gauss.

The global spectra of radio pulsars are fairly complicated. Figure 6.44 displays energy spectra for a handful of radio pulsars.

For older pulsars, there is a clear evidence for soft X-ray excess probably associated with the surface of the neutron star. In the case of the Crab pulsar, this thermal emission is not yet visible.

X-Ray Observations of Radio Pulsars

Geminga is a radio-quiet pulsar with period $P = 0.237$ s, surface magnetic field $B_* \simeq 1.6 \times 10^{12}$ G, and characteristic age $\tau = P/2\dot{P} = 3.4 \times 10^5$ yr. The ROSAT PSPC spectrum of Geminga has been fitted with a two-component model, either

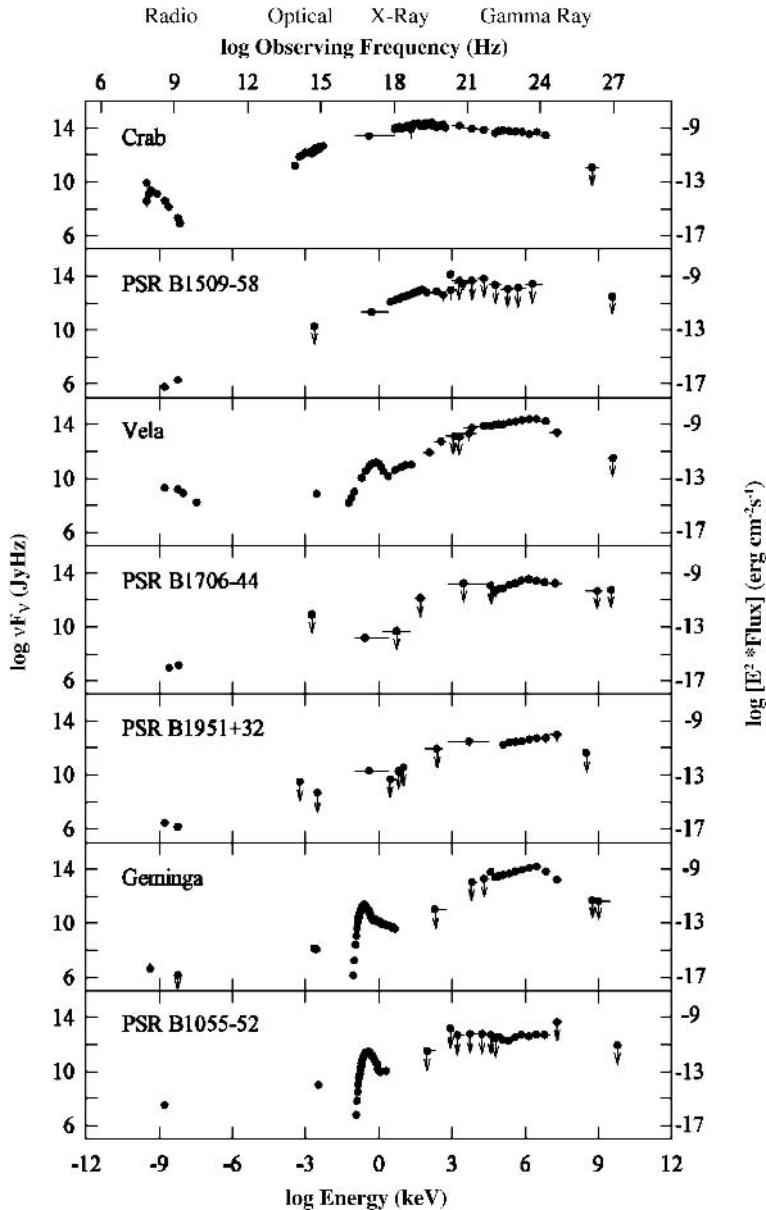


Fig. 6.44. The global spectra of seven radio pulsars sorted with age increasing from *top* to *bottom*. Figure adapted from Becker and Pavlov [66]

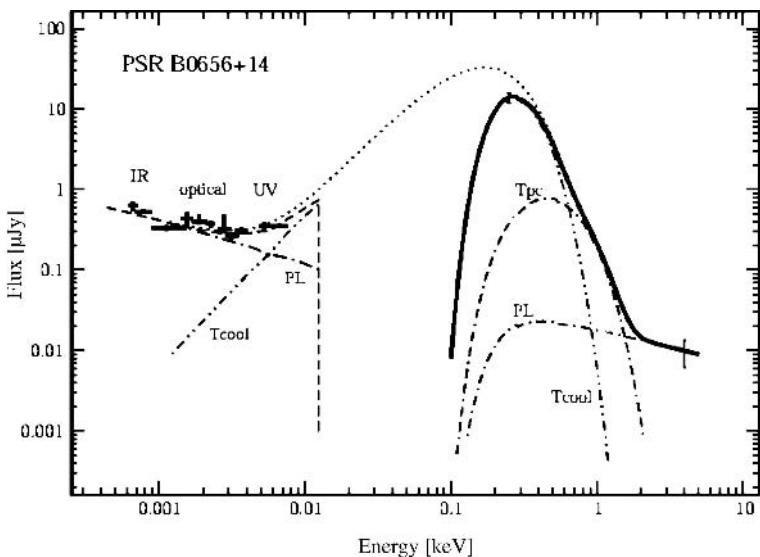


Fig. 6.45. The observed X-ray spectrum of PSR B0656. The intrinsic BB spectrum (dotted line) is strongly absorbed by the ISM

two black-bodies or a black-body plus a power law. Both fits gave a satisfactory description of the spectrum. This ambiguity was resolved by combining ROSAT and ASCA data; the bulk of the soft X-ray flux is a black-body, and the hard X-ray spectrum has a power-law profile. The soft X-ray component is parameterized as a black-body with $T = 5.77 \times 10^5$ K and a bolometric luminosity 1.47×10^{31} ergs/s; the hard X-ray spectrum is fitted by a power law with energy index 0.47 and luminosity 8.13×10^{29} ergs/s. The luminosities were calculated from the parallax distance of 160 pc from the HST measurements.

PSR 1055–52 and PSR 0656+14 resemble Geminga in many respects: period, loss rate of rotational energy, and characteristic age (see Table 6.9). They also have been fitted with a two-component model. The soft X-ray component has generally been interpreted as thermal emission from all or a part of the surface of the neutron stars. However, the source of the harder X-ray component has not been determined. ROSAT data from PSR 1055–52 were fitted to a black-body together with a power law of energy index 0.4–0.5. Although black-body models to describe the hard X-ray component for a joint fit of ROSAT and ASCA data have been favored, a reexamination of the same ASCA data shows that a power law with energy index 0.5 ± 0.3 and luminosity 1.5×10^{30} ergs s⁻¹ (calculated at $d = 500$ pc) fits the spectrum equally well. PSR 1055–52 possesses a large pulsed fraction of 0.73 ± 0.33 at higher X-ray energies.

At present, about 20 pulsars have been detected in soft X-rays, but few of them have sufficient photons for comprehensive spectral and temporal analysis. The above five objects, observed in the X-ray energy range 0.1–10.0 keV, represent pulsars of various classes. They all have periods of about 0.3 s. Geminga, PSR 1055–52, and

PSR 0656+14 have characteristic ages $\tau \simeq 10^5$ yr, whereas PSR 1929+10 and PSR 0950+08 have $\tau \simeq 10^7$ yr. Strong pulsed gamma-ray emission is detected from Geminga and PSR 1055–52, but such radiation is still uncertain for PSR 0656+14. While Geminga, PSR 1055–55, and PSR 0656+14 have typical dipole component magnetic fields $B_* \simeq 10^{12}$ G, PSR 1929+10 and PSR 0950+08 have weaker fields ($B_* < 10^{12}$ G). Luminosities have a great uncertainty associated with the measurements of distance. Thermal soft X-ray emission is observed from Geminga, PSR 1055–52, and PSR 0656+14, and the emitting area is comparable to the entire surface of a neutron star. A thermal hard X-ray component is observed from PSR 1929+10, PSR 0950+08, and PSR 0656+14 from an area that occupies only a small fraction of the neutron star surface. A power-law spectrum in hard X-rays is observed from the gamma-ray pulsars Geminga and PSR 1055–52. Light-curves of the hard X-rays of both gamma-ray pulsars have pulsed fractions exceeding 50%, which could be an indicator of nonthermal emission.

Gamma-Ray Pulsars

The Energetic Gamma-Ray Experiment Telescope (EGRET) on the Compton Gamma Ray Observatory (CGRO) has detected six pulsars with strong gamma-ray emission above 100 MeV: Crab, Vela, Geminga, PSR 1706–44, PSR 1055–52, and PSR 1951+32. Similarities among the gamma-ray spectra and most of the light-curves suggest that a similar powerful accelerator is operating in each of their magnetospheres.

Models for magnetospheric accelerators fall into two main classes [184] (see also Fig. 6.41). For a polar-cap accelerator generally assumed to power the radio emission of a pulsar, the acceleration of primary particles takes place relatively near the surface of a neutron star. In the central dipole approximation for the stellar magnetic field, all the radiation and pair production resulting from primaries that have passed through such accelerators would be restricted to be within the open field-line bundle that links the polar cap to the light cylinder ($R_L = c/\Omega$). Outer-magnetospheric accelerators have been proposed for the emission of the energetic photons from gamma-ray pulsars. These accelerators could achieve much higher potential drops along magnetic field lines

$$\Delta V \simeq \frac{\Omega^2 B_* R^3}{c^2} \simeq 10^{14} \text{ volt}, \quad (6.320)$$

with B_* the surface dipole field, than can polar cap accelerators. Pair production within such accelerators (e.g. from accelerator-produced GeV gamma-rays colliding with keV X-rays from the stellar surface or its neighborhood) will result in a net flow of primary e^\pm out of the star being balanced by an inward flow of e^\pm from the starward end of the accelerator. Because of extensive e^\pm flows generated by both polar cap and outer-magnetospheric accelerators, it is difficult to see how both could survive on the same field lines, pairs from one gap would be expected to quench the other gap. However, a polar cap might be linked to both kinds of accelerator as long as the field

lines through each are different (e.g. upward curving with a net positive [negative] charge near the light cylinder for one and downward curving negative [positive] for the other, or lines passing through the charge reversing surface $\Omega \cdot \mathbf{B} = 0$ for the one but not for the other).

Arguments for important outer-magnetospheric accelerators as well as polar cap ones in gamma-ray pulsars but not in ordinary ones, include the following: (i) apparently different death lines for the two families; for example, ordinary pulsar radio emission appears to cease when polar cap accelerator e^\pm production is no longer possible while gamma-ray pulsar emission seems to be quenched when putative outer-magnetospheric accelerators would no longer be able to sustain the mechanism of collision between GeV gamma-rays and keV X-rays (thus Geminga and PSR 1055–52 could maintain outer-magnetospheric accelerators, but PSR 1929+10 and PSR 0950+08 could not; the former are gamma-ray pulsars, while the latter are not); (ii) strong optical emission coincident with X-ray emission from the Crab pulsar that is difficult to explain unless its source is synchrotron radiation in the outer magnetosphere; (iii) escape of 10 GeV gamma-rays from PSR 1951+32 and PSR 1706–44, which seems more difficult to accomplish, if these gamma-rays originate in the neighborhood of strong polar cap fields. From an outer-magnetospheric accelerator, because almost the full potential drop $\Delta V \simeq 10^{14}$ V will be radiated away by each electron in the accelerator, we indeed expect just such a luminosity,

$$L_\gamma \simeq \dot{N}_e e \Delta V \simeq 10^{34} \text{ erg/s}, \quad (6.321)$$

for \dot{N}_e the maximum current through the accelerator (Goldreich–Julian current)

$$\dot{N}_e \simeq \frac{\Omega^2 B_* R^3}{ec} \simeq 10^{32} \text{ s}^{-1}. \quad (6.322)$$

The spectrum of the Crab pulsar in the ROSAT and ASCA regime is dominated by the radiation in two subpulses. That radiation is almost certainly from the same region and particles that are the source of its high-energy (up to at least 1 GeV) gamma-rays, but not pairs created on closed field lines in the near magnetosphere discussed above. For X-rays from the latter we should look at the radiation between these subpulses, which should also have a broader angular spread than that in the two subpulses. The spectrum of this interpulse radiation has been isolated. A power law with an energy index of 0.5 (smaller by 0.32 than that in the subpulses) is natural, consistent with that for Geminga and PSR 1055–52. This radiation extends up to 200 keV.

Relativistic inflowing particles radiate away much of their energy before reaching the polar cap. With constant bombardment of these particles at a rate of 10^{32} s^{-1} (the Goldreich–Julian current), the polar cap would radiate X-rays. However, no evidence for such a strongly heated polar cap is found in Geminga or PSR 1055–52. The total thermal X-ray luminosity of PSR 1055–52 is 2.3×10^{32} ergs/s, whereas for Geminga it is 1.5×10^{31} ergs/s. In both cases, X-rays seem to be emitted from a large fraction of the surface of the neutron star. We attribute the absence of observed hot polar caps to the formation of a reflective e^\pm “blanket” above the stellar surface.

What Can we Learn from High-Energy Emission?

The broad-band spectra contain a lot of information on the emission processes in the magnetosphere of a pulsar. There exist a number of magnetospheric emission models in the literature, but no theory reached a widespread acceptance. The physics of the magnetosphere of rotating neutron stars is still one of the big mysteries.

What is known is that the radio emission of pulsars is a coherent process. On the other hand, the optical, X- and gamma-ray emission must be incoherent. Therefore, the fluxes in these energy bands are directly proportional to the densities of the radiating particles, no matter which radiation process (synchrotron, curvature radiation or inverse Compton emission) is at work in the given energy. Various processes have been advanced for the high-energy emission:

- Nonthermal emission from charged particles accelerated in the pulsar magnetosphere. The energy distribution of these particles is in general a power law, and the emission would then also be characterized by power-law-like spectra. The energy of the photons would extend into the TeV band.
- X-ray and gamma-ray emission from the pulsar wind.

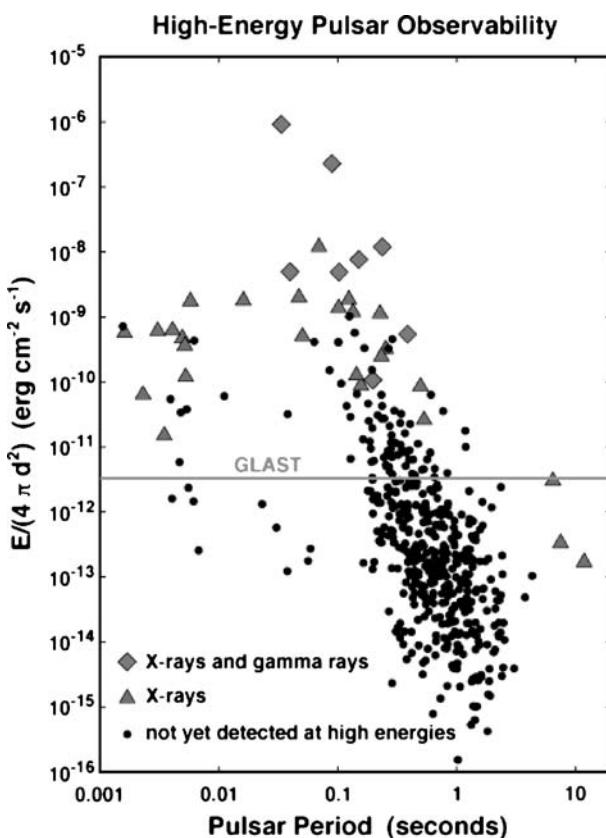


Fig. 6.46. Pulsars detected so far in high-energy emission, compared to the sensitivity of GLAST. The spin-down energy loss is plotted as a function of the spin period of the neutron star

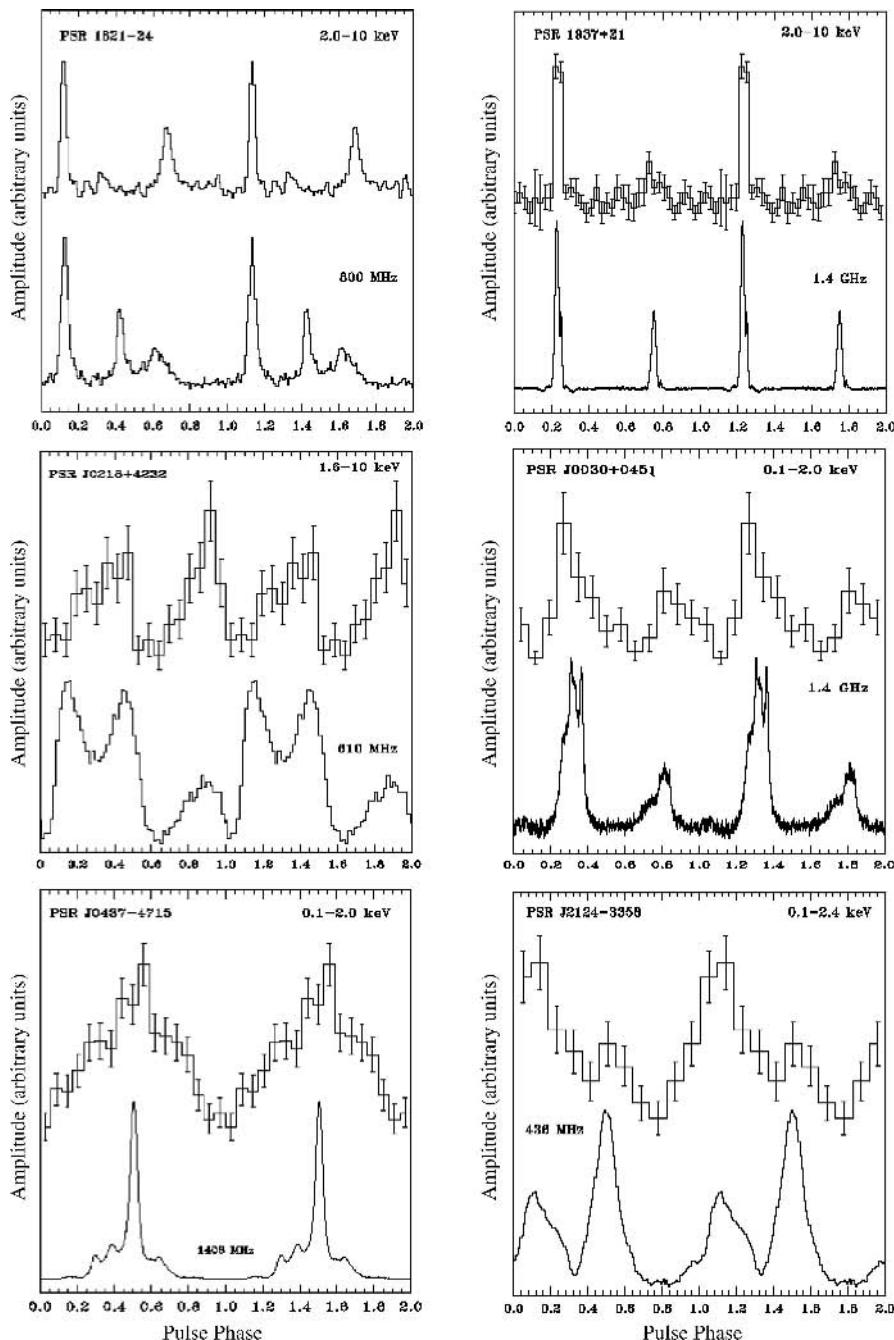


Fig. 6.47. The observed radio (*lower curve*) and X-ray light-curves of millisecond pulsars. In most cases, both polar caps are visible and the X-ray emission appears to be modulated in the same way as the radio emission

- Photospheric emission from the hot surface of a cooling neutron star. A modified black-body spectrum would result in this case with low-amplitude intensity variations with the rotational period (Fig. 6.47). This emission would occur in the optical and soft X-ray region (see Fig. 6.45).
- Thermal soft X-ray emission from the neutron star's polar cap which are heated by relativistic particles, or by Joule heating due to currents crossing the polar caps. This would lead to a strong modulation of the X-ray flux, especially for millisecond pulsars (Fig. 6.47).

The Gamma-ray Large Area Space Telescope (GLAST) will open this high-energy world to exploration and help us to answer these questions. With GLAST, the sensitivity range for the detection of radio pulsars in the gamma-ray range can be extended towards weaker sources (Fig. 6.46).

6.6.4 Accretion-Powered Neutron Stars and the Mass–Radius Relation

Over 90% of the strong Galactic X-ray sources appear to fall into two distinct groups: (i) the high-mass X-ray binaries (HMXBs) and (ii) the low-mass X-ray binaries (LMXBs). These two groups differ in a number of physical characteristics (see Table 6.11). Binary pulsars and single millisecond pulsars are the descendants of the X-ray binaries containing an accreting neutron star.

High-Mass X-Ray Binaries – HMXBs

There are about 130 known HMXBs and 25 have well-measured orbital parameters. There are ~ 40 pulsating HMXB sources with typical pulse periods between 10–300 seconds (the entire observed range spans between 0.069 seconds and 20 minutes).

Table 6.11. The two main classes of strong Galactic X-ray sources

	HMXB	LMXB
X-ray spectra:	$kT \geq 15$ keV (hard)	$kT \leq 10$ keV (soft)
Type of time variability:	regular X-ray pulsations no X-ray bursts	only a very few pulsars often X-ray bursts
Accretion process:	wind (or atmos. RLO)	Roche-lobe overflow
Timescale of accretion:	10^5 yr	10^7 – 10^9 yr
Accreting compact star:	high B -field NS (or BH)	low B -field NS (or BH)
Spatial distribution:	Galactic plane	Galactic center and spread around the plane
Stellar population:	young, age $< 10^7$ yr	old, age $> 10^9$ yr
Companion stars:	luminous, $L_{\text{opt}}/L_x > 1$ early-type O(B) stars $> 10 M_{\odot}$ (Pop. I)	faint, $L_{\text{opt}}/L_x \ll 0.1$ blue optical counterparts $\leq 1 M_{\odot}$ (Pop. I and II)

Among the systems with $P_{\text{orb}} \leq 10$ days and $e \leq 0.1$ are the strong sources and ‘standard’ systems such as Cen X-3 and SMC X-1. These are characterized by the occurrence of regular X-ray eclipses and double-wave ellipsoidal light variations produced by tidally deformed (‘pear-shaped’) giant or subgiant companion stars with masses $> 10 M_{\odot}$. However, the optical luminosities ($L_{\text{opt}} > 10^5 L_{\odot}$) and spectral types of the companions indicate original ZAMS masses $\geq 20 M_{\odot}$, corresponding to O-type progenitors.

The companions have radii $10\text{--}30 R_{\odot}$ and (almost) fill their critical Roche lobes. In a number of pulsating sources, such as X0115+63 (and Her X-1, an intermediate-mass X-ray binary system) there are absorption/emission features in the X-ray spectrum which are most probably cyclotron lines, resulting from magnetic fields with strengths $B \simeq 5 \times 10^{12}$ G. Among the standard HMXBs, there are at least two systems that are thought to harbor black holes: Cyg X-1 and LMC X-3.

Another group of HMXBs consists of the moderately wide, eccentric binaries with $P_{\text{orb}} \simeq 20\text{--}100$ days and $e \simeq 0.3\text{--}0.5$. A new third (sub-)group has recently been proposed. These systems have $P_{\text{orb}} \simeq 30\text{--}250$ days and small eccentricities $e \leq 0.2$. Together these two groups form a separate subclass of HMXBs: the Be-star X-ray binaries. In the Be-star X-ray binaries the companions are rapidly rotating B-emission stars situated on, or close to, the main sequence (luminosity class III–V). There are more than 50 such systems known making them the most numerous class of HMXBs. The Be stars are deep inside their Roche lobes, as is indicated by their generally long orbital periods (≥ 15 days) and by the absence of X-ray eclipses and of ellipsoidal light variations. According to the luminosities and spectral types, the companion stars have masses in the range about $8\text{--}20 M_{\odot}$ (spectral types O9–B3, III–V). The X-ray emission from the Be-star X-ray systems tends to be extremely variable, ranging from complete absence to giant transient outbursts lasting weeks to months (Fig. 6.48). During such an outburst episode one often observes orbital modulation of the X-ray emission, due to the motion of the neutron star in an eccentric orbit.

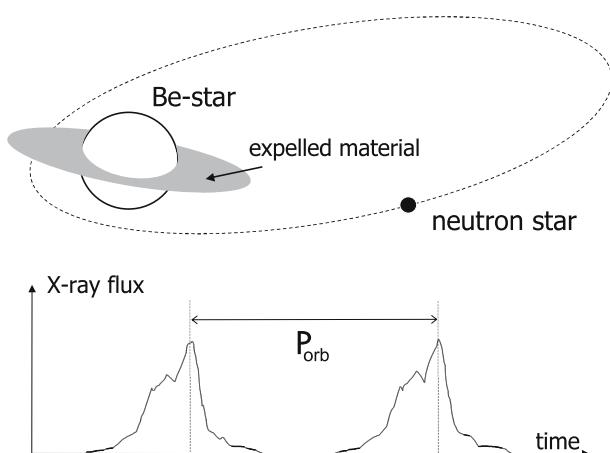


Fig. 6.48. Schematic model of a Be-star X-ray binary system. The neutron star moves in an eccentric orbit around the Be star which is not filling its Roche lobe. However, near the periastron passage the neutron star accretes circumstellar matter, ejected from the rotating Be star, resulting in an X-ray burst lasting several days

the irregular optical outbursts generally observed in Be stars, which indicate sudden outbursts of mass ejection, presumably generated by rotation-driven instability in the equatorial regions of these stars. While the Be-star X-ray binaries are transient sources (often unobservable for months to years) the ‘standard’ systems are persistent X-ray sources. HMXBs are located along the Galactic plane among their OB-type progenitor stars⁶.

Low-Mass X-Ray Binaries – LMXBs

Orbital periods have been measured for some 30 of these systems. They range from 11 min. to 17 days, similar to the orbital periods of cataclysmic variables. Only in the few widest of these systems one can observe the spectrum of the optical companion. In all other systems, the optical spectrum is that of the hot accretion disk. The LMXBs are very seldom X-ray pulsars. The reason is their relatively weak magnetic fields $10^9 \sim 10^{11}$ G, which is expected to result from accretion-induced field decay. On the other hand these sources show X-ray bursts (sudden thermonuclear fusion of accreted matter at the surface of the neutron star, which are suppressed if the magnetic field strength is $> 10^{11}$ G. For this reason such bursts are not observed in HMXBs.

The discovery of (kilohertz) quasiperiodic oscillations (QPOs) in the X-ray flux of LMXBs has provided a clear timing signature of the accreting neutron stars and black holes in these systems. In the past decade much insight of detailed accretion physics and testing of general theory of relativity has been revealed by observations and “beat frequency” models. There is more than a dozen of LMXBs systems for which there is strong evidence for the presence of a black hole (see Sect. 6.5).

Low-mass X-ray binaries (LMXB) are binary stars where one of the components is a black hole or a neutron star. The other component (donor) is filling its Roche lobe and therefore transfers mass to the compact star. The donor is less massive than the compact object, and usually on the main sequence. In some cases the donor can be either degenerate or evolved, (subgiant or red giant). About one hundred LMXB’s

⁶ Besides the HMXB and LMXB systems with companion stars $> 10 M_\odot$ and $\leq 1 M_\odot$, respectively, there must exist also a large number of Galactic compact binaries with companion star masses in the interval $1\text{--}10 M_\odot$. These are the so-called *intermediate-mass X-ray binaries* (IMXBs). IMXBs have recently been recognized as a class of their own. IMXB systems are not easily observed as a result of a simple selection effect against X-ray sources with intermediate-mass companions. The reason is the following: “Standard” HMXBs have evolved (sub-)giant companions which are massive enough to have a strong stellar wind mass-loss rate (typically $\dot{M}_{\text{wind}} \simeq 10^{-6} M_\odot \text{ yr}^{-1}$) sufficient to power a bright X-ray source, via an accreting neutron star or black hole, for $10^5\text{--}10^6$ yr. The LMXBs often evolve slowly on a nuclear time-scale ($10^8 \sim 10^9$ yr) and the majority of the transferred material is usually funnelled onto the compact object, via an accretion disk, yielding accretion rates of $10^{-10}\text{--}10^{-8} M_\odot \text{ yr}^{-1}$. In IMXBs the companions are not massive enough to produce sufficiently high wind mass-loss rates to power an observable X-ray source. Her X-1 and Cyg X-2 are systems of this type. In the latter system the companion presently has a mass $< 1 M_\odot$, but it is highly overluminous for this mass.

have been detected from the Milky Way. Several LMXB have been discovered from globular clusters.

A typical LMXB emits almost all of its radiation in X-rays, and typically less than one percent in visible light, so they are among the brightest objects in the X-ray sky, but relatively faint in visible light. The apparent magnitude is typically around 20 magnitudes. The brightest part of the system is the accretion disk around the compact object. The orbital periods of LMXBs range from a few hours to a few days.

X-Ray Bursters and Rotation of Neutron Stars

Many accreting neutron stars erupt in spectacular thermonuclear conflagrations every few hours to days. These events are known as Type I X-ray bursts, or simply X-ray bursts. Powerful new X-ray observatories, the Rossi X-ray Timing Explorer (RXTE), the Italian–Dutch BeppoSAX mission, XMM–Newton and Chandra have enabled the discovery of entirely new phenomena associated with thermonuclear burning on neutron stars. Some of these new findings include: (i) the discovery of millisecond (300–600 Hz) oscillations during bursts, so-called “burst oscillations,” (ii) a new regime of nuclear burning on neutron stars which manifests itself through the generation of hours long flares about once a decade, now referred to as “superbursts,” (iii) discoveries of bursts from low accretion rate neutron stars, and (iv) new evidence for discrete spectral features from bursting neutron stars.

It is perhaps surprising that nuclear physics plays such a prominent role in the phenomenology of an accreting neutron star, as the gravitational energy released per accreted baryon (of mass m_p) is $GMm_p/R \approx 200$ MeV is so much larger than the nuclear energy released by fusion (≈ 5 MeV when a solar mix goes to heavy elements). Indeed, if the accreted fuel was burned at the rate of accretion, any evidence of nuclear physics would be swamped by the light from released gravitational energy. The only way the nuclear energy can be seen is when the fuel is stored for a long period and then burns rapidly (as in Type I bursts and superbursts).

The successful association of thermonuclear instabilities with X-ray bursts made a nice picture of a recurrent cycle that consists of fuel accumulation for several hours to days followed by a thermonuclear runaway that burns the fuel in ~ 10 –100 seconds. It also secured the identification of the accreting objects as neutron stars. The mass donors – the ultimate source of the thermonuclear fuel – are typically old, Population II objects or in some cases, degenerate helium or perhaps carbon/oxygen white dwarfs. The accreted composition is important, as the nuclear ashes and burst properties depend on the accreted mix of light elements. Unfortunately, in most cases, we have little information on the composition of the accreted fuel. Of the approximately 160 known LMXB’s about 70 are observed to produce bursts [380].

Early theoretical studies noted the likely importance of spreading of the thermonuclear burning front around the neutron star surface. Because nuclear fuel is burned in a time much shorter than it takes to accrete a critical pile it is unlikely

Table 6.12. X-ray burst oscillation sources and properties

Source	Frequency (Hz)	$\Delta\nu_{\text{QPO}}$	$P_{\text{orb}}(\text{hr})$
4U 1728–34	363	280–363	?
4U 1636–53	581	250–320	3.8
KS 1731–260	524	260	?
Galactic center	589	?	?
Aql X-1	549	?	19.0
4U 1702–429	330	315–344	?
MXB 1658–298	567	?	7.1
4U 1916–053	270	290–348	0.83
4U 1608–52	619	225–325	?
SAX J1808.4–3658	401	?	2.0
SAX J1750.8–2980	601	?	?

that ignition conditions will be achieved over the entire surface simultaneously. It appears more likely that burning is initiated locally and then spreads laterally, eventually engulfing all fuel-loaded parts of the neutron star. For conditions most prevalent in burst sources, the front may spread via convective deflagration, at lateral speeds of up to $\approx 5 \times 10^6 \text{ cm s}^{-1}$. Such speeds can account for the subsecond rise times of some bursts, but the time required for burning to engulf the entire star is still long compared to the spin periods of accreting LMXB neutron stars (milliseconds). Moreover, if the burning front is not strongly convective, then a patchy distribution of nuclear fuel is possible. These considerations suggest that during bursts the rotation of the neutron star can modulate the inhomogeneous or localized burning regions, perhaps allowing for direct observation of the spin of the neutron star.

An exciting development in the past decade has been the discovery of high frequency (300–600 Hz) X-ray brightness oscillations during bursts. These modulations are now commonly called *burst oscillations*. They were first discovered with the PCA onboard RXTE in bursts from the LMXB 4U 1728–34. As of this writing, burst oscillation detections have been claimed for an additional 10 sources, one of which is the 401 Hz accreting millisecond pulsar SAX J1808.4–3658, whose spin period is precisely known. Table 6.12 provides a catalog of the known burst oscillation sources and summarizes some of their salient properties.

The oscillation frequency during a burst is typically not constant. Most commonly the frequency increases as the burst progresses, that is, the evolution can be characterized as a chirp (Fig. 6.49). Most bursts show spin-up toward some limiting, or asymptotic frequency, however, there are exceptions to this rule of thumb. For example, 4U 1636–53 has shown a spin down in the cooling tail. This burst also showed an unusually long thermal tail which may have been related to a reheating episode having some connection to the spin-down. Spin downs are apparently rare. The observed frequency drifts are generally $< 1\%$ of the mean frequency, and bursts which have detectable pulsations during the rising phase show the largest frequency

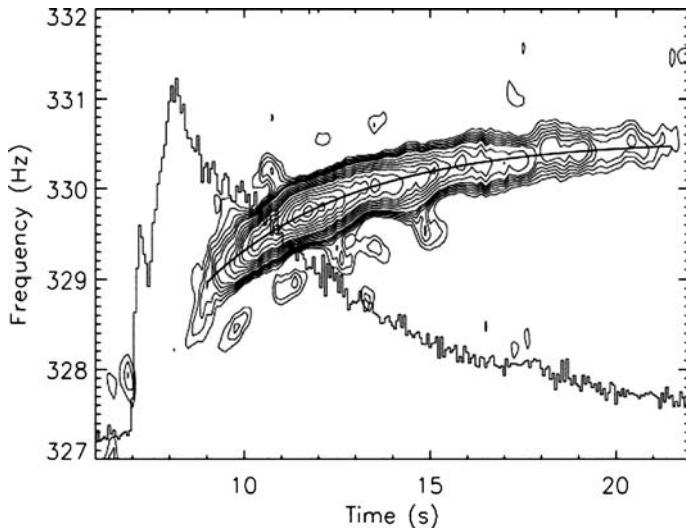


Fig. 6.49. X-ray burst profile and millisecond pulsations (*contour lines*) in the LMXB 4U 1702–429. Shown are contours of constant power spectral density as a function of frequency and time. The *solid curve* shows the best fitting exponential model. During the outburst, the rotation of the neutron star becomes visible. Figure adapted from [379]

shifts. This indicates that the process responsible for the frequency evolution begins with the start of a burst, and not when oscillations are first detected within a burst.

Strohmayer and Markwardt [379] studied the frequency evolution in bursts from 4U 1702–429 and 4U 1728–34 (Fig. 6.49). They found the frequency in these bursts could be modelled as a smooth exponential recovery of the form, $v(t) = v_0(1 - \delta_v e^{-t/\tau})$, where v_0 , δ_v , and τ are the asymptotic frequency, the fractional frequency drift, and the recovery timescale, respectively. With this form they were able to recover coherent signals, with coherence values, $Q \equiv v_0/\Delta v_0 > 4500$ in some bursts. Figure 6.49 shows an example of a burst from 4U 1702–429 with exponential frequency evolution. These results support the existence of a reference frame on the neutron star, perhaps the nuclear burning layer, in which the oscillations are coherent or nearly so. This frame, however, cannot be rigidly connected to the bulk of the neutron star, because the torque required to change the spin frequency of the star by $\approx 1\%$ in only 10 seconds is unphysically large. This implies the existence of shearing in the surface layers of the neutron star. In the exponential model, the total amount of phase shearing is simply $\phi_{\text{shear}} = v_0\delta_v(1 - e^{-T/\tau})$, where T is the length of the pulse train. For typical bursts this value ranges from about 4–8, suggesting that the burning layer “slips” this many revolutions over the underlying neutron star during the duration of the pulsations. The amount of phase shearing has implications for the surface magnetic field strength, as a sufficiently strong field will enforce corotation.

QPOs in Accreting Neutron Stars

The fastest periodic signals seen to date in astronomy have been observed by RXTE in accreting neutron stars. In February 1996, two groups analyzing data at Goddard Space Flight Center began to see evidence for these ultrafast variations in the X-ray light-curves of 4U 1728–34 (led by Tod Strohmayer of Goddard) and Sco X-1 (led by Michiel van der Klis of the University of Amsterdam). The oscillations they observed are nearly periodic and have frequencies of about 1000 hertz – hence the name *kilohertz (kHz) quasiperiodic oscillations (QPO)*.

To date these signals have been seen in about 20 neutron stars in the Galaxy. Figure 6.50 shows a power spectrum of Sco X-1, in which the relative amplitude of oscillation as a function of frequency is plotted. One can see two tiny peaks near 1000 Hz which represent the kHz QPO. They are dwarfed by the much larger peak near 10 Hz showing the usual QPO, which has been known and studied in many systems for over 10 years. Since the amount of power associated with an oscillation varies as the frequency times the amplitude of the power spectrum at that point, there is far more power in the kHz QPO than the normal QPO. In many of the sources, one sees two frequencies which vary with flux – the exact relationship is different for different objects. For some of the sources the difference between the

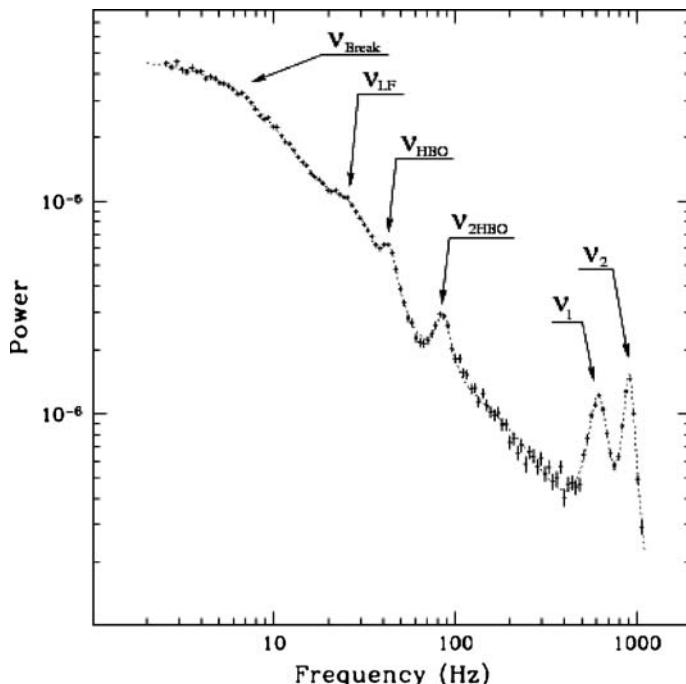


Fig. 6.50. The power spectrum for the X-ray source Sco X-1. The total observing time is about 10,000 seconds. Various QPOs are visible: HBOs and their second harmonic, as well as ν_1 and ν_2 , and the break frequency ν_{Break}

frequencies remains constant, and is thought to be telling us the rotation rate of the neutron star. The rotation rates seen so far are all between about 250 and 350 hertz (i.e. rotations per second). The difference $v_2 - v_1$ is nearly constant at 300–500 Hz. In the ultracompact binary system 4U 1820–30 it has been found that above a certain flux level, the frequencies saturate to a maximum. The current understanding of this saturation is that we are seeing the innermost stable circular orbit (ISCO) around the neutron star, a radius predicted by general relativity inside of which matter must inexorably spiral down to smaller radii. Besides the resonance peaks, the power spectrum can generally be fitted by a broken power law

$$P(v) = A v^{-\alpha} \left[1 + (v/v_{\text{Break}})^{\beta} \right]^{-1} \quad (6.323)$$

with $\alpha \simeq 1$ and $\beta \simeq 1$. Beyond the break frequency, the spectrum steepens considerably.

The kHz QPO are the most important scientific result to date of RXTE, based on the high timing capability of the satellite. One of the most developed model to account for kHz QPO is the *sonic point model* of Cole Miller, Fred Lamb, and Dimitrios Psaltis [294]. The model follows the detailed trajectories of blobs of gas leaving the inner edge of the accretion disk around a neutron star and spiraling down to the neutron star surface, taking into account the effects of general relativity.

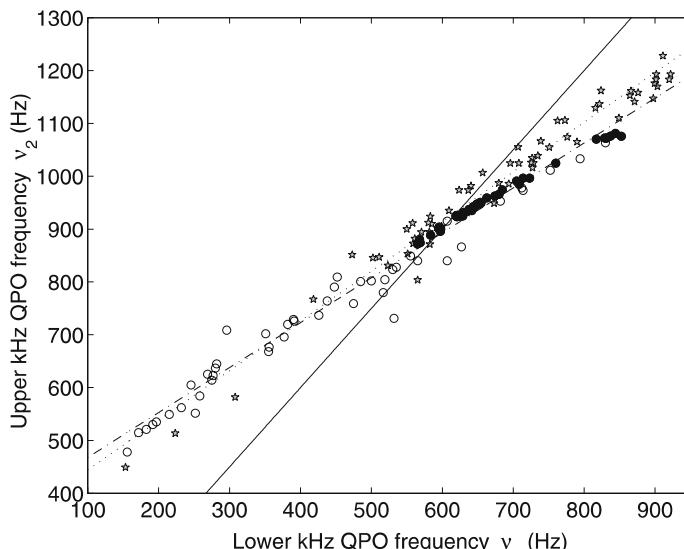


Fig. 6.51. Plot of upper kHz QPO frequency (v_2) vs. lower kHz QPO frequency (v_1) for the Atoll and Z sources for which double kHz QPOs have been published. Empty circles correspond to Z sources, stars to Atoll sources, and filled circles to the Sco X-1 data. The *dotted line* is the best linear fit to the Atoll points, the *dot-dashed line* is the best linear fit to the Z points (excluding Sco X-1), and the *thick line* represents a fixed 3:2 ratio. Errors of the frequencies are typically around a few Hz. Figure adapted from Belloni et al. [69]

The gas spirals inward supersonically and collides with the surface of the neutron star. These authors utilize observations of the highest frequency kHz QPO to place constraints on the allowed equation of state for neutron stars. This relationship between pressure and density for a gas made up of dense neutrons determines the mass-radius relation for neutron stars. There is currently no consensus as to the origin of these QPOs, nor on what physical parameters determine their frequencies, which have been identified with various characteristic frequencies in the inner accretion flow.

The upper QPO frequency ν_2 is associated in most models with the Keplerian frequency at the inner edge of the disk

$$r_T \simeq 15 \text{ km} \frac{M}{M_\odot} \left(\frac{\nu_2}{1000 \text{ Hz}} \right)^{-2/3}. \quad (6.324)$$

This radius r_T is then a function of the mass accretion rate in order to explain the frequency variations. The truncation of a Keplerian disk beyond the innermost stable orbit can have various reasons: truncation by a weak magnetosphere of the neutron star, or by a transition to a hotter optically thin inner disk (see Sect. 10.3). In modern theories of accretion towards compact objects, the geometrically thin optically thick disk always gets truncated at a certain radius. The lower QPO frequency ν_1 is usually explained in terms of a beat frequency between ν_2 and the rotation frequency ν_* of the neutron star surface, $\nu_1 = \nu_2 - \nu_*$. This would require rotational frequencies in the range of 300–500 Hz, as observed in many sources (see Table 6.12). The origin of the low-frequency ν_{HBO} is less clear, but interestingly it is close to the frequency of the Lense–Thirring precession frequency of the inner disk around the rotational axis of the rotating neutron star [301, 374]

$$\nu_{\text{LT}} = \frac{8\pi^2 I_*}{Mc^2} \nu_K^2 \nu_* \simeq 13.2 \text{ Hz} \frac{I_{45}}{M/M_\odot} \nu_{K,3}^2 \nu_{*,2.5}. \quad (6.325)$$

General relativity predicts that a rotating body with moment of inertia I_* produces a frame-dragging (or Lense–Thirring) effect: the orbital plane of a test particle in a nonequatorial orbit precesses about the body's symmetry axis.

Belloni et al. [69] have analyzed all published frequencies, ν_1 and ν_2 , of the twin kilohertz quasiperiodic oscillations (kHz QPOs) in bright neutron star low-mass X-ray binaries. The two frequencies are well correlated but, contrary to recent suggestions, the frequency–frequency correlation is significantly different from a $\nu_2 = (3/2)\nu_1$ relation. The QPO frequencies are checked whether they cluster around a region where $\nu_2/\nu_1 \approx 3/2$. These results do not provide any useful information about a possible underlying resonance mechanism in Sco X-1. The (unbiased) distribution of QPO frequencies is consistent with a uniform distribution at a 2.4σ level. In general, the distribution of the kHz QPO frequencies is not uniform and has multiple peaks, which have no analogy in the distribution of points in the spectral color–color diagrams of these sources.

Alternatively, one can constrain the compactness of the neutron star by measuring the amplitudes of the periodic oscillations (which give direct evidence for the neutron

star spin rates) during thermonuclear bursts. General relativity predicts that signals arising from near the surface of the neutron star should be deflected by the strong gravitational field. The more compact a star is, the more that strongly varying signals are homogenized into smoothly varying signals. This occurs because the light from any localized spot on the surface of the star would be bent considerably, so that the spot might be seen even when it was on the far side of the star from our line-of-sight. The observed amplitudes of the $\simeq 300$ Hz oscillations tell us that the lower limit on the radius for a neutron star is about 7.7 kilometers (for an assumed mass of 1.4 solar masses). A neutron star could not be more compact and yet show oscillation amplitudes as large as those seen.

Typical neutron star flares last 10 seconds, a thousand times shorter and 500–1000 times less energetic than this flare in binary star system 4U 1820–30, observed with the Rossi X-ray Timing Explorer. The three-hour flare from the tiny neutron star released 100 times more energy than the Sun does in a year.

The neutron star in 4U 1820–30 is in tight orbit with a low-mass dwarf star burning mostly helium instead of hydrogen. Their orbital period of 11 minutes is the shortest known of any binary system. The neutron star's strong gravitational potential pulls gas away from the companion star. This gas crashes down on the surface of a neutron star. When enough gas builds up on the neutron star surface – in this case, helium gas – the increased pressure raises the temperature and initiates helium fusion, a nuclear reaction that manifests itself as an X-ray flare. X-ray flares often erupt on neutron stars in binary systems several times a day.

Problems

6.1. Proton–Neutron Ratio: Show that the proton–neutron ratio for neutron star matter is close to 1/8, $n_p \simeq n_n/8$.

Show that the Fermi momentum of the electrons in neutron star matter largely exceeds the Fermi momentum of the free beta decay.

6.2. TOV Solution for Free Neutron Gas: Integrate the TOV equations for a free neutron gas with EoS $P = K\rho^\Gamma$ and determine the maximal mass, maximal density and minimal radius. Make a plot with mass vs. central density.

6.3. TOV Solver: Develop a solver for the TOV equations which can handle equations of state in numerical form. Apply this solver to various equations of state and determine the mass as a function of central density.

6.4. Maximal Core Masses: The mass of a neutron star can be divided into the mass of the envelope, where EoS is known, and the mass M_c of the core, where the EoS is unknown (say above saturation density). Relativity requires that

$$M_c < R_c/2 \quad (6.326)$$

and therefore for given central density

$$M_c < \frac{1}{2} \sqrt{\frac{3}{8\pi\varrho_c}} , \quad R_c < \sqrt{\frac{3}{8\pi\varrho_c}} . \quad (6.327)$$

In order to study the possible ranges, we investigate the redshift factor $\alpha = \exp \Phi(r) > 0$. α has its minimal value at the center. Derive the following equation for the redshift factor from the TOV equations

$$\sqrt{1 - \frac{2M(r)}{r}} \frac{1}{r} \frac{d}{dr} \left[\sqrt{1 - \frac{2M(r)}{r}} \frac{1}{r} \frac{d\Phi}{dr} \right] = \frac{\alpha}{r} \frac{d}{dr} \left(\frac{M(r)}{r} \right) . \quad (6.328)$$

Use this equation to show the inequality

$$\sqrt{1 - \frac{2M(r)}{r}} \frac{1}{r} \frac{d\Phi}{dr} \leq \left[\int_0^r dr \frac{r}{\sqrt{1 - 2M(r)/r}} \right]^{-1} . \quad (6.329)$$

Derive from this equation the inequality

$$M_c \leq \frac{2}{9} R_c \left[1 - 6\pi R_c^2 P_c + \sqrt{1 + 6\pi R_c^2 P_c} \right] . \quad (6.330)$$

6.5. Post-Newtonian Expansion: Calculate the Christoffel symbols for the post-Newtonian expansion of the metric and derive the equations of motion for point masses (neutron stars and black holes).

6.6. Binary Orbit Decay: Calculate the first, second and third time derivative of the moment of inertia for a binary orbit

$$I_{ik} = \int \varrho_0(\mathbf{x}) x^i x^k d^3x . \quad (6.331)$$

6.7. Pulsar Diagram: Use the data on the web interface:

www.atnf.csiro.au/research/pulsar/psrcat to compile a (P, \dot{P}) pulsar diagram.

6.8. Eddington Luminosity: What is the Eddington luminosity for a neutron star?

6.9. Crab Nebula: Compile all data about the spectrum for the Crab Nebula from the literature (radio, infrared, optical, UV, X-rays, gamma-rays and TeV region). Make with these data a spectrum of the Crab Nebula, $\log v$ vs. $\log[v F_v]$, where the dimension of the spectral energy vF_v is in units of Jy Hz. Derive from this spectrum the total luminosity of the Crab Nebula and compare this with the energy loss of the Crab pulsar.

6.10. Neutrino and Photon Transport in Neutron Stars: Discuss the diffusion approximation for the photon and neutrino transport in neutron stars. Use the momentum formulation for the Boltzmann equation to recover the neutrino (photon) transport equations for the number density N_ν , the number flux F_ν and the number

source term S_N , together with the mean neutrino energy density J_ν , energy flux H_ν , pressure P_ν and energy source term S_E

$$\frac{\partial[N_\nu/n_B]}{\partial t} + \frac{\partial[4\pi r^2 \exp \Phi F_\nu]}{\partial r} = \exp \Phi \frac{S_N}{n_B} \quad (6.332)$$

$$\begin{aligned} \frac{\partial[J_\nu/n_B]}{\partial t} + P_\nu \frac{\partial[1/n_B]}{\partial t} \\ + \exp(-\Phi) \frac{\partial[4\pi r^2 \exp(2\Phi) H_\nu]}{\partial r} = \exp \Phi \frac{S_E}{n_B}. \end{aligned} \quad (6.333)$$

6.11. Masses of Neutron Stars: Most of the binary pulsars have white dwarf stars as companions, and the binary systems are practically circular. Find out a method based on pulsar timing to measure the masses for such systems.

6.12. Self-Gravitating Bose–Einstein Condensates: The most general metric for a time-dependent spherically symmetric object can be written as

$$ds^2 = -\alpha^2 dt^2 + a^2(dr + \beta dt)^2 + b^2 r^2 d\Omega^2, \quad (6.334)$$

where $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$ is the metric on the unit two-sphere, β is the r -component of the shift vector, and α , β , a and b are functions of t and r only. The corresponding extrinsic curvature has only two nonvanishing components

$$K_j^i = \text{diag}(K_r^r, K_\theta^\theta, K_\phi^\phi). \quad (6.335)$$

Together with the four metric functions, we have to deal with six out of 16 possible geometrical variables.

- (i) Write down the inverse metric in matrix form.
- (ii) Calculate all connection forms and Ricci tensors, necessary for the 3+1 split of Einstein's equations.
- (iii) Show that the source terms in Einstein's equations are given by

$$\varrho = \frac{|\Phi|^2 + |\Pi|^2}{2a^2} + \frac{m^2|\phi|^2}{2} \quad (6.336)$$

$$S_r = -\frac{\Pi^* \Phi + \Pi \Phi^*}{2a} \quad (6.337)$$

$$S_r^r = \varrho - m^2|\phi|^2 \quad (6.338)$$

$$S_\theta^\theta = \frac{|\Pi|^2 - |\Phi|^2}{2a^2} - \frac{m^2|\phi|^2}{2} \quad (6.339)$$

$$S_\phi^\phi = S_\theta^\theta \quad (6.340)$$

$$S = \frac{3|\Pi|^2 - |\Phi|^2}{2a^2} - \frac{3m^2|\phi|^2}{2}. \quad (6.341)$$

Here, we have defined the following quantities

$$\Phi = \phi', \quad \Pi = \frac{a}{\alpha}(\dot{\phi} - \beta\phi'). \quad (6.342)$$

This is an example, where the pressure is not isotropic, $P_r \neq P_\perp$.

(iv) Calculate the Hamiltonian and momentum constraints for this system, as well as the four evolution equation for \dot{a} , \dot{b} , \dot{K}_r^r , and \dot{K}_θ^θ . Use the Klein–Gordon equation to show that this results in equations for ϕ , $\dot{\Phi}$ and $\dot{\Pi}$.

Fix the coordinate system by the maximal slicing condition, $\text{Tr}(K) = 0$, which fixes the lapse function (see Sect. 2.8). Given maximal slicing, the remaining spatial coordinate freedom is fixed by the isotropic condition

$$a = b = \psi^2(t, r), \quad (6.343)$$

such that the spatial metric is conformally flat

$$ds_{(3)}^2 = \psi^4(t, r) (dr^2 + r^2 d\Omega^2). \quad (6.344)$$

(v) Derive the constraints equation for the lapse function α following from the condition $\text{Tr}(K) = 0$.

(vi) Formulate the remaining evolution equations for $\psi(t, r)$, K_r^r , and $\phi(t, r)$.

7 Rapidly Rotating Neutron Stars

The Schwarzschild solution of general relativity is static and spherically symmetric. As such, this solution can only describe nonrotating stellar structures. In this section we derive a more general geometry that is able to cover rotating objects which are axisymmetric. This class of gravitational fields contains several types of astrophysical objects:

- rapidly rotating neutron stars (as numerical solutions of Einstein’s equations);
- rotating black holes (the Kerr solution as an analytic solution of Einstein’s equations), will be discussed in Chap. 8;
- star–toroid systems (e.g. a neutron star surrounded by a self-gravitating torus);
- self-gravitating disk solutions;
- toroidal black hole solutions.

The calculation of the structure of rapidly rotating relativistic stars has a long tradition in astrophysics. It dates back to the pioneering work of Carter [106], Bardeen [55], and Bardeen and Wagoner [57]. Successful numerical methods have been advanced in the last 15 years by Komatsu et al. [235] (KEH), Cook, Shapiro and Teukolsky [123, 124], Bonazzola et al. [83, 84], and many others since then (see Stergioulas [376]). The Kerr solution as an analytic solution of Einstein’s equations will be discussed in the next section. Here, we concentrate on general properties of spacetimes for rotating objects. In particular, we derive the Ricci tensors for such geometries and discuss the numerical solutions of Einstein’s equations for rotating neutron stars. Finally, the question of the stability for such objects is a central issue of the present research.

There are two main effects that distinguish a rotating relativistic star from its nonrotating counterpart: The shape of the star is flattened by centrifugal forces (an effect that first appears at second order in the rotation rate), and the local inertial frames are dragged by the rotation of the source of the gravitational field. While the former effect is also present in the Newtonian limit, the latter is a purely relativistic effect. The study of the dragging of inertial frames in the spacetime of a rotating star is assisted by the introduction of the local Zero Angular-Momentum Observers (ZAMO). These are observers whose worldlines are normal to the hypersurfaces $t = \text{const}$, they are also called Eulerian observers. Then, the metric function ω is the angular velocity of the local ZAMO with respect to an observer at rest at infinity. Also, the redshift factor α

is the time dilation factor between the proper time of the local ZAMO and coordinate time t (which is the proper time at infinity) along a radial coordinate line.

7.1 Spacetime of Stationary and Axisymmetric Rotating Bodies

We first consider the metric appropriate to stationary and axisymmetric spacetimes. For this purpose it is convenient to take as two coordinates the time $t(=x^0)$ and the azimuthal angle $\phi(=x^1)$ about the axis of symmetry. The stationary and axisymmetric character of the spacetime requires that the metric coefficients be independent of t and ϕ , i.e.

$$g_{\alpha\beta} = g_{\alpha\beta}(x^2, x^3), \quad (7.1)$$

where the coordinates x^2 and x^3 are the two remaining coordinates.

Besides stationarity and axisymmetry, we also require that the spacetime is invariant against simultaneous inversion of time t and angle ϕ , i.e. to the transformation $t \rightarrow -t$ and $\phi \rightarrow -\phi$. The physical meaning of this condition is that the source of the gravitational field has motions that are pure rotational about the axis of symmetry, since the energy-momentum tensor as the source of the metric will have the same symmetry. Differential rotation is, however, allowed. This invariance will require that

$$g_{t2} = 0 = g_{t3} = g_{\phi 2} = g_{\phi 3}. \quad (7.2)$$

Under these assumptions, the metric will have the following form

$$\begin{aligned} ds^2 &= g_{tt} dt^2 + 2g_{t\phi} dt d\phi + g_{\phi\phi} d\phi^2 \\ &\quad + [g_{22} (dx^2)^2 + 2g_{23} dx^2 dx^3 + g_{33} (dx^3)^2]. \end{aligned} \quad (7.3)$$

All the metric functions are only functions of x^2 and x^3 . A further reduction is possible as a consequence of the fact that a metric of a two-space, (x^2, x^3) , represented by the brackets $[\dots]$ with positive or negative signature can always be brought to the diagonal form

$$ds_2^2 = \pm \exp(2\mu) [(dx^2)^2 + (dx^3)^2] \quad (7.4)$$

by a mere coordinate transformation with $\exp(2\mu)$ as a function of x^2 and x^3 . For the proof of this Lemma one has simply to consider a particular coordinate transformation $x'^2 = \Phi(x^2, x^3)$ and $x'^3 = \Psi(x^2, x^3)$. One can then show that

$$\Phi_{,2} = \sqrt{g} [g^{32}\Psi_{,2} + g^{33}\Psi_{,3}] \quad (7.5)$$

$$\Phi_{,3} = -\sqrt{g} [g^{32}\Psi_{,2} + g^{33}\Psi_{,3}] \quad (7.6)$$

together with the integrability condition

$$[\sqrt{g} g^{AB} \Psi_{,A}]_{,B} = 0, \quad A, B = 2, 3 \quad (7.7)$$

will satisfy the requirements. As a consequence of this Lemma we can reduce the number of free metric coefficients to four. For many purposes in the following, it is however useful to consider only a diagonal form for the two-space so that

$$\begin{aligned} ds^2 = & -\exp(2\nu) dt^2 + \exp(2\psi) (d\phi - \omega dt)^2 \\ & + \exp(2\mu_2) (dx^2)^2 + \exp(2\mu_3) (dx^3)^2. \end{aligned} \quad (7.8)$$

The five functions ν, ψ, ω, μ_2 and μ_3 are now functions of x^2 and x^3 . It is important to note here that in this form we left one gauge freedom for the coordinate choice. It is this gauge freedom which is responsible for different expressions of the metric found in the literature. The so-called quasi-isotropic gauge corresponds to $\gamma_{r\theta} = 0$ and $\gamma_{\theta\theta} = r^2 \gamma_{rr}$, if we choose the remaining coordinates to be of spherical type, $x^2 = r$ and $x^3 = \theta$.

7.1.1 Physical Interpretation of the Metric

It is clear that the two functions μ_2 and μ_3 are of geometric origin and describe the curvature of the two-space (x^2, x^3) . As a consequence of the coordinate choice, the function $\exp \psi$ has the meaning of a cylindrical radius. So there remain two essential potentials which represent the gravitational field of the rotating body. In analogy to the Schwarzschild solution, the first function $\exp \nu$ represents the classical gravitational field. It is often written as $\alpha \equiv \exp \nu$, denoting the lapse function (or redshift factor) for the time-slicing of a spacetime.

Bardeen Observers

In order to understand the meaning of ω we consider an orthonormal tetrad frame \mathbf{e}_a for $a = 0, 1, 2, 3$ (called a Bardeen observer, locally nonrotating observer, or zero angular-momentum observer ZAMO)

$$e_0 = \exp(-\nu) (\partial_t + \omega \partial_\phi) \quad (7.9)$$

$$e_1 = \exp(-\psi) \partial_\phi \quad (7.10)$$

$$e_2 = \exp(-\mu_2) \partial_2 \quad (7.11)$$

$$e_3 = \exp(-\mu_3) \partial_3. \quad (7.12)$$

One can then easily show that $g(e_a, e_b) = \eta_{ab}$ satisfies the Minkowski conditions. The corresponding one-form basis is then

$$\Theta^0 = \exp \nu dt \quad (7.13)$$

$$\Theta^1 = \exp \psi (d\phi - \omega dt) \quad (7.14)$$

$$\Theta^2 = \exp \mu_2 dx^2 \quad (7.15)$$

$$\Theta^3 = \exp \mu_3 dx^3. \quad (7.16)$$

This shows that the metric can be written as

$$ds^2 = \eta_{ab} \Theta^a \Theta^b. \quad (7.17)$$

In this form the metric is locally Minkowskian, and the \mathbf{e}_a represent locally an inertial frame. But this is true for any other observer obtained by means of a local Lorentz transformation, $\mathbf{e}'_a = \Lambda_a^b \mathbf{e}_b$.

Motion and Frame-Dragging

Now let us consider a velocity field u at each point

$$u^t = \frac{dt}{ds}, \quad u^\phi = \frac{d\phi}{ds} = \Omega u^t, \quad u^A = \frac{dx^A}{ds} = u^t v^A \quad (7.18)$$

with proper time ds and

$$\Omega = \frac{d\phi}{dt} \quad (7.19)$$

as the angular velocity with respect to time t . The components

$$v^A = \frac{dx^A}{dt}, \quad A = 2, 3 \quad (7.20)$$

denote the poloidal velocity. The normalization of the velocity field provides the relation

$$1 = [\exp 2\nu - \exp 2\psi(\Omega - \omega)^2 - \exp 2\mu_2(v^2)^2 - \exp 2\mu_3(v^3)^2] (u^t)^2 \quad (7.21)$$

or

$$u^t = \frac{\exp(-\nu)}{\sqrt{1 - V^2}} = \frac{1}{\alpha \sqrt{1 - V^2}} \quad (7.22)$$

with the norm

$$V^2 = \exp(-2\nu) [\exp(2\psi)(\Omega - \omega)^2 + \exp(2\mu_2)(v^2)^2 + \exp(2\mu_3)(v^3)^2]. \quad (7.23)$$

This is usually written for the Lorentz factor W with respect to ZAMO as

$W = \alpha u^t.$

(7.24)

With this velocity field we can associate the following components with respect to the local observer

$$u^{(a)} = \Theta_\mu^a u^\mu \quad (7.25)$$

or explicitly

$$u^{(0)} = W = \frac{1}{\sqrt{1 - V^2}} \quad (7.26)$$

$$u^{(1)} = \frac{\exp(\psi)(\Omega - \omega)}{\alpha\sqrt{1 - V^2}} = \frac{\exp(\psi)(\Omega - \omega)}{\sqrt{\alpha^2 - \bar{V}^2}} \quad (7.27)$$

$$u^{(A)} = \frac{\exp(\mu_A)v^A}{\alpha\sqrt{1 - V^2}} = \frac{\exp(\mu_A)v^A}{\sqrt{\alpha^2 - \bar{V}^2}}. \quad (7.28)$$

Therefore, a particle moving on a circular orbit with angular velocity Ω will be assigned a rotational velocity $\exp \psi(\Omega - \omega)/\alpha$ with respect to ZAMOs. This particle can be considered to be at rest in this system, whenever $\Omega = \omega$, i.e. with respect to fixed stars the particle will still orbit around the central source with angular velocity ω . This effect is called the **dragging of inertial frames**. In the following we will show that the source for this potential ω is the angular momentum J , in the sense that far away from the source $\omega \propto J/r^3$. This potential drops therefore very rapidly with increasing radius, and frame-dragging effects can only be observed in the immediate vicinity of the surface of compact objects (see, however, also next section).

Induced Differential Rotation

The above result can also be discussed in another way. Let us write the velocity field in the form

$$u = u^t(k + \Omega m) + u^A \partial_A \quad (7.29)$$

with its angular velocity Ω and poloidal velocity u^A . $\lambda = u_\phi/E = j/E$ denotes its specific angular momentum which is conserved in geodetic motions, $E = u^\alpha k_\alpha = -u_t$ is the energy measured at infinity. Then Ω can be solved from the definition

$$\Omega = \frac{u^\phi}{u^t} = \frac{g^{\phi\phi}u_\phi + g^{t\phi}u_t}{g^{tt}u_t + g^{t\phi}u_\phi} = \frac{g^{\phi\phi}(u_\phi/u_t) + g^{t\phi}}{g^{tt} + g^{t\phi}(u_\phi/u_t)} \quad (7.30)$$

resulting in the expression

$$\Omega = \Omega(\lambda) = \omega + \frac{\alpha^2}{R^2} \frac{\lambda}{1 - \lambda\omega}$$

(7.31)

with the cylindrical radius $R \equiv \exp \psi = r \sin \theta \exp \bar{\psi}$. Here we use the inverse metric in the form

$$g^{tt} = -\frac{1}{\alpha^2} \quad (7.32)$$

$$g^{t\phi} = -\frac{\omega}{\alpha^2} \quad (7.33)$$

$$g^{\phi\phi} = \frac{\alpha^2 - R^2\omega^2}{\alpha^2 R^2}. \quad (7.34)$$

This means that the angular velocity is the sum of two terms: the classical one given by the specific angular momentum and the frame-dragging ω from the rotation of absolute space. Near the horizon of a black hole, for example, where $\alpha \rightarrow 0$, the angular velocity of matter is completely dominated by the frame-dragging effect. Whatever the specific angular momentum of the incoming matter is, this matter is forced to rotate with the local angular velocity ω . When matter falls, e.g. into a nonrotating black hole, it is forced to zero rotation near the horizon despite its angular momentum. This behavior can never be mimicked in a kind of pseudo-Newtonian accretion, since the effect of frame-dragging is absent in the Newtonian world.

7.1.2 Geodetic and Lense–Thirring Precession

An effect predicted by general relativity and also known as **frame-dragging**, in which the orbit of a small body orbiting around a rotating massive one is slightly perturbed by the rotation. The effect was first predicted by Austrian physicists Joseph Lense and Hans Thirring in 1918.

A rotating massive body drags space and time around with it. A gyroscope orbiting Earth tends to tilt away from the plane of its orbit because the Earth is dragging it. In curved spacetime, a gyroscope orbiting a body will precess

$$\frac{d\mathbf{S}}{d\tau} = \boldsymbol{\Omega} \times \mathbf{S} \quad (7.35)$$

with the geodetic precession

$$\boldsymbol{\Omega}_{\text{GP}} = -\frac{3}{2c^2} \mathbf{v} \times \nabla \Phi. \quad (7.36)$$

$\Phi = -GM/r$ is the external gravitational potential of the body. For this case, we find for the geodetic precession

$$\boldsymbol{\Omega}_{\text{GP}} = \frac{3}{2c^2} \frac{GM}{|\mathbf{x}|^3} (\mathbf{v} \times \mathbf{x}). \quad (7.37)$$

A spin \mathbf{J} of the body produces in addition the Lense–Thirring precession

$$\boldsymbol{\Omega}_{\text{LT}} = -\frac{G}{c^2 |\mathbf{x}|^3} \left(\mathbf{J} - \frac{3(\mathbf{J} \cdot \mathbf{x})\mathbf{x}}{|\mathbf{x}|^2} \right). \quad (7.38)$$

Gravity Probe B is an experiment developed by NASA and Stanford University to precisely test these two predictions of Albert Einstein's general theory of relativity, measuring how space and time are warped by the presence of the Earth and how the Earth's rotation drags spacetime around with it. For a gyroscope orbiting near the Earth, the geodetic effect leads to a tilting of the gyroscope's spin axis in the plane of the orbit (Fig. 7.2). This effect is predicted by general relativity theory to be 150

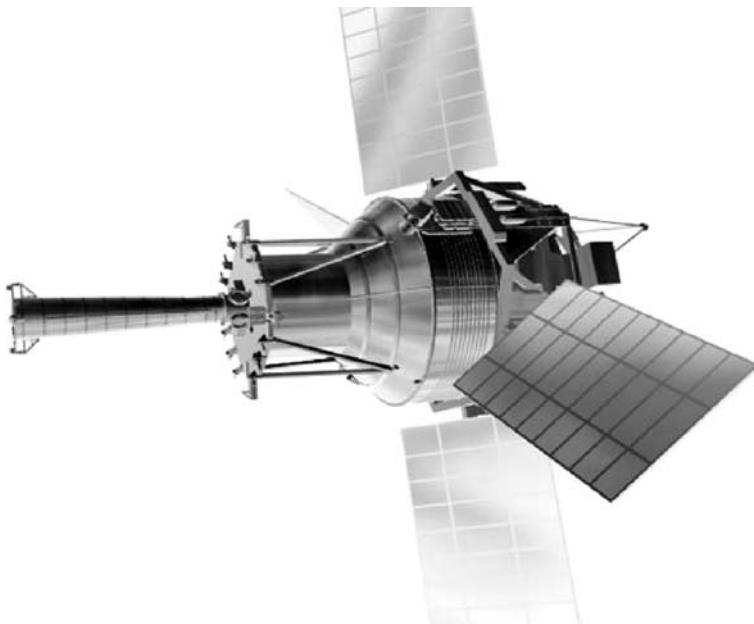


Fig. 7.1. The satellite Gravity Probe B will measure the geodetic precession and the Lense–Thirring effect in a satellite orbiting the Earth. Credits: Gravity Probe B (JPL)

times larger than the frame-dragging (Lense–Thirring) effect. GP-B will measure this effect to a part in 10,000.

At the heart of the experiment inside Gravity Probe B are four gyroscopes (Fig. 7.1). The gyroscopes in Gravity Probe B are not flywheels, but electrically supported spheres spinning in a vacuum. At the center of each of the four gyroscopes is a jewel-like sphere of fused quartz the size of a ping-pong ball. According to NASA, the ultrasmooth spheres, coated with niobium, are the roundest objects ever made by man. The GP-B gyros, which are performing perfectly in orbit, will be listed in the forthcoming edition of the *Guinness Book of World Records* as being the roundest objects ever manufactured. Recent measurements show that the actual characteristic spin-down period of the GP-B gyros exceeds 10,000 years – well beyond the requirement of 2300 years. The spheres are enclosed in chambers to prevent disruption from sound waves. They are chilled to near absolute zero to prevent their molecular structures from creating disturbances. The gyroscopes are 30 million times more accurate than any gyroscope ever built. The gyroscopes spin at 10,000 revolutions a minute. Very sensitive magnetometers detect any changes in the gyroscope's spin axis. To measure each orbit closely, a tracking telescope is used to align the gyroscopes with a guide star. A magnetic-field measuring device (SQUID) records changes relative to the guide star. The gyro readout measurements from the SQUID magnetometers have unprecedented precision, detecting fields to 10^{-13} gauss, less than one trillionth of the strength of Earth's magnetic field. The

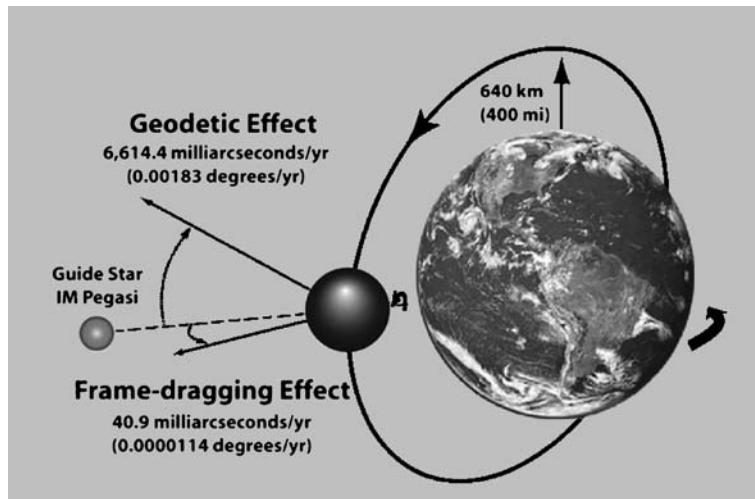


Fig. 7.2. The satellite Gravity Probe B will measure the geodetic precession and the Lense Thirring effect in a satellite orbiting the Earth. Credits: Gravity Probe B (JPL)

satellite was launched to a 640-km high orbit on April 20, 2004, on a Delta 2 rocket from Vandenberg Air Force Base, on the central coast of California. Gravity Probe B spacecraft circles the globe every 97.5 minutes, crossing over both poles. Over the course of a year, the anticipated spin axis drift for the geodetic effect is a minuscule angle of 6614.4 milliarcseconds, and the anticipated spin axis drift for the frame-dragging effect is even smaller, only 40.9 milliarcseconds. The GP-B program will not release the scientific results obtained during the mission until after the science phase has concluded. It is critically important to thoroughly analyze the data to ensure its accuracy and integrity prior to releasing the results.

Previously, an international team of NASA and university researchers has found the first direct evidence of the phenomenon that the Earth is dragging space and time around itself as it rotates. Researchers believe they have detected the effect by precisely measuring shifts in the orbits of two Earth-orbiting laser-ranging satellites, the Laser Geodynamics Satellite I (LAGEOS I), a NASA spacecraft, and LAGEOS II, a joint NASA/Italian Space Agency (ASI) spacecraft. LAGEOS II, launched in 1992, and its predecessor, LAGEOS I, launched in 1976, are passive satellites dedicated exclusively to laser ranging, which involves sending laser pulses to the satellite from ranging stations on Earth and then recording the round-trip travel time. The LAGEOS satellites are passive vehicles covered with retroreflectors designed to reflect laser beams transmitted from ground stations. Given the well-known value for the speed of light, this measurement enables scientists to determine precisely the distances between laser ranging stations on Earth and the satellite. LAGEOS is designed primarily to provide a reference point for experiments that monitor the motion of the Earth's crust, measure and understand the wobble in the Earth's axis of rotation, and collect information on the Earth's size, shape, and gravitational field.

Both satellites are spherical bodies with an aluminum shell wrapped around a brass core. The design was a compromise between numerous factors, including the need to be as heavy as possible to minimize the effects of nongravitational forces vs. being light enough to be placed in a high orbit and the need to accommodate as many retroreflectors as possible vs. the need to minimize surface area to minimize the effects of solar pressure. The materials were chosen to reduce the effects of the Earth's magnetic field on the satellite's orbit. 426 cube-corner retroreflectors are imbedded in the satellites' surface. 422 of these are made of fused silica glass while the other four are made of germanium. The vehicles have no on-board sensors or electronics, and are not attitude controlled. Science is performed by reflecting laser light from the vehicle's 426 retroreflectors. The design life is 50 years.

The Lense–Thirring effect, or frame-dragging effect, represents the manifestation that mass-currents, i.e. the angular momentum of the central mass, produce on a moving test particle a force analogous to the Lorentz force of classical electromagnetism on a moving charged particle. This non-Newtonian gravitomagnetic force is produced by the gravitomagnetic field of the rotating mass, a new feature of Einstein's theory. Considering a satellite orbiting the Earth, the Lense–Thirring effect is responsible of a secular shift of the satellite ascending node longitude, Ω , as well as of its argument of perigee, ω . A satellite undergoes a perigee precession due to the Earth's angular momentum J_E

$$\Delta\Omega_{\text{LT}} = \frac{2GJ_E}{c^2a^3(1-e^2)^{3/2}} \simeq 3 \text{ arcsec/century} \quad (7.39)$$

with a corresponding secular drift of the node

$$\Delta\Omega_{\text{LT}} = -\frac{6GJ_E \cos i}{c^2a^3(1-e^2)^{3/2}}. \quad (7.40)$$

The effect is very small, about 2 m/yr on the node of the LAGEOS-type satellites, the best tracked Earth's satellites through the powerful Satellite Laser Ranging (SLR) technique. The SLR technique allows the determination of the position of both LAGEOS satellites with precision of a few mm in their normal points, and a root-mean-square of their range residuals of about 2–3 cm over a 15 days time span. Therefore, from an analysis of the orbit of both LAGEOS satellites it is possible to measure the relativistic precession predicted by Einstein's theory.

7.1.3 On General 3+1 Split of Spacetime

The form of the metric (7.8) is a special realization of the so-called 3+1 split of spacetime (see Sect. 2.8). In a deep way, general relativity is a theory of four-dimensional spacetime. But humans experience spacetime as a three-dimensional space that evolves in time. For the purpose of simulating spacetime on a computer, it is also natural to *slice* four-dimensional spacetime into a sequence of three-dimensional spaces, labelled by time (Fig. 2.7). In the computer, one begins with initial data on

one such slice, then evolves them for a short moment of time to get the next slice, and then discards the first slice and moves on to the next slice. Spacetime is built up slice by slice.

In the 3+1 formalism spacetime is decomposed into a one parameter family of space-like slices. This is the most intuitive approach as it corresponds to splitting spacetime into space and time. The field equations then comprise of four constraint equations that have to be satisfied on the first slice and are then satisfied throughout spacetime by virtue of the contracted Bianchi identities, six evolution equations, and four gauge variables that represent the gauge freedom of general relativity and can be used to simplify the equations. The 3+1 formalism is the most commonly used approach to numerical relativity. In particular, it is very useful for studies of regions with matter. However, the formulation suffers drawbacks in the description of radiation, because bringing in null infinity to finite coordinates via conformal compactification cannot be done in a straightforward manner. At the present time, most simulations based on the 3+1 formalism do not incorporate realistic boundary conditions at outer end of the computational grid.

All geometrical objects are now split into components with respect to this time-slice of spacetime. In particular, the components of the energy-momentum tensor T are expressed in terms of the following components

$$E = T(n, n) = T_{\mu\nu} n^\mu n^\nu \quad (7.41)$$

$$J_\alpha = -\gamma_\alpha^\mu T_{\mu\nu} n^\nu \quad (7.42)$$

$$S_{\alpha\beta} = \gamma_\alpha^\mu \gamma_\beta^\nu T_{\mu\nu} \quad (7.43)$$

$$S = S_\alpha^\alpha. \quad (7.44)$$

E is now the energy density measured by an adapted observer of four-velocity n^α (often called an Eulerian observer), \mathbf{J} the momentum flow and \mathbf{S}_{ij} the corresponding stress tensor. The entire stress-energy tensor can therefore be written as

$$T^{\alpha\beta} = E n^\alpha n^\beta + n^\alpha J^\beta + J^\alpha n^\beta + S^{\alpha\beta}. \quad (7.45)$$

Similarly, the Ricci and Einstein tensor have to be split according to these time-slices, as discussed in Sect. 2.8.

For a perfect fluid with stress-energy tensor

$$T^{\mu\nu} = (\varrho + P) u^\mu u^\nu + P g^{\mu\nu}, \quad (7.46)$$

where u^α is the four-velocity of the fluid, ϱ the energy density and P the pressure, respectively, we define the Lorentz factor to be

$$W = -n_\mu u^\nu = \alpha u^t. \quad (7.47)$$

From this we get the energy density

$$E = W^2(\varrho + P) - P \quad (7.48)$$

and the momentum density

$$J^i = (E + P)v^i, \quad (7.49)$$

where the fluid three-velocity v^i is related to the spatial components of the fluid four-velocity by

$$u^i = W \left(v^i - \frac{\beta^i}{\alpha} \right). \quad (7.50)$$

As usual, the Lorentz factor can be expressed as $W = 1/\sqrt{1 - v^2}$, where $v = \sqrt{v^i v_i}$ is the physical fluid velocity measured by the Eulerian observer. Finally, the stress tensor is given by

$$S_{ij} = (E + P)v_i v_j + P \gamma_{ij}. \quad (7.51)$$

In a rotating configuration, the fluid four-velocity is just given by

$$u = u^t \left(\frac{\partial}{\partial t} + \Omega \frac{\partial}{\partial \phi} \right), \quad (7.52)$$

where $\Omega = u^\phi/u^t$ is the fluid angular velocity, as seen by an inertial observer at rest at infinity (fixed stars).

7.2 Einstein's Field Equations for Rotating Objects

In this section we discuss the split of the Einstein tensor for the spacetimes of rotating objects. This is a special case of the decomposition of Einstein's field equations within the ADM formalism obtained in Sect. 2.8, here given as a decomposition of the Ricci tensor

$$R_{00} = \frac{1}{\alpha} \bar{\Delta} \alpha + \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \text{Tr}(\mathbf{K}) - \text{Tr}(\mathbf{K}^2) \quad (7.53)$$

$$R_{0i} = D_i \text{Tr}(\mathbf{K}) - (D \cdot \mathbf{K})_i \quad (7.54)$$

$$\begin{aligned} \mathbf{Ric}(g) &= \mathbf{Ric}(\gamma) - 2\mathbf{K}^2 + \text{Tr}(\mathbf{K})\mathbf{K} \\ &\quad - \frac{1}{\alpha} (D \otimes D)\alpha - \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \mathbf{K}. \end{aligned} \quad (7.55)$$

γ is the metric of the hypersurface with $\mathbf{Ric}(\gamma)$ as its Ricci tensor and \bar{R} its Ricci scalar, \mathbf{K} is the symmetric matrix formed by the extrinsic curvature

$$2\alpha K_{ij} = -\frac{\partial \gamma_{ij}}{\partial t} + D_i \beta_j + D_j \beta_i. \quad (7.56)$$

D is the covariant derivative for the hypersurface. When the first equation is coupled to the energy-momentum tensor, we arrive at the generalized Poisson equation

$$\frac{1}{\alpha} \bar{\Delta} \alpha + \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \text{Tr}(\mathbf{K}) - \text{Tr}(\mathbf{K}^2) = 4\pi G [2E - \epsilon + 3P]. \quad (7.57)$$

Here we used the trace of the energy-momentum tensor, $T = -\epsilon + 3P$, for a perfect fluid and the energy density measured by the ZAMO

$$E = \gamma^2(\epsilon + PV^2). \quad (7.58)$$

This equation shows that the Poisson equation for the redshift factor is satisfied in the Newtonian limit, where $\alpha \simeq 1 + \Phi$ and all nonlinear terms are neglected. It also shows that the extrinsic curvature leads to various corrections for the Poisson equation. Since in our case, the gravitational fields are stationary, all time derivatives vanish. In addition, we only get two constraints equations, since S_ϕ is the only nonvanishing momentum density of the rotating matter.

7.2.1 Ricci Tensors of Time-Slices

In the above decomposition of Einstein's equations we need the Ricci tensors $\mathbf{Ric}(\gamma)$ of the time-slices (for the derivation, see Appendix D), or in components $\bar{R}_{ij} = \bar{Q}_i^m(e_m, e_j)$, with the following six components

$$\bar{R}_{11} = -\frac{1}{R} \nabla_2[R\Psi_2] - \Psi_3[\nabla_3\mu_2] - \frac{1}{R} \nabla_3[R\Psi_3] - \Psi_2[\nabla_2\mu_3] \quad (7.59)$$

$$\begin{aligned} \bar{R}_{22} = & -\frac{1}{R} \nabla_2[R\Psi_2] - \Psi_3[\nabla_3\mu_2] \\ & - \exp(-\mu_2) \nabla_3[\exp(\mu_2) \nabla_3\mu_2] - \exp(-\mu_3) \nabla_2[\exp(\mu_3) \nabla_2\mu_3] \end{aligned} \quad (7.60)$$

$$\begin{aligned} \bar{R}_{33} = & -\exp(-\mu_2) \nabla_3[\exp(\mu_2) \nabla_3\mu_2] \\ & - \exp(-\mu_3) \nabla_2[\exp(\mu_3) \nabla_2\mu_3] \end{aligned} \quad (7.61)$$

$$\bar{R}_{12} = 0 \quad (7.62)$$

$$\bar{R}_{13} = 0 \quad (7.63)$$

$$\bar{R}_{23} = -\frac{1}{R} \nabla_3[R\Psi_2] + \Psi_3[\nabla_2\mu_3]. \quad (7.64)$$

By summation we get the Ricci scalar on the hypersurface

$$\begin{aligned} \bar{R} = & -2 \left[\frac{1}{R} \nabla_2[R\Psi_2] + \Psi_3(\nabla_3\mu_2) + \frac{1}{R} \nabla_3[R\Psi_3] + \Psi_2(\nabla_2\mu_3) \right] \\ & - 2\Delta(\mu_2, \mu_3), \end{aligned} \quad (7.65)$$

where we have defined the second-order elliptic operator

$$\begin{aligned} \Delta(\mu_2, \mu_3) = & \exp(-\mu_2) \nabla_3[\exp(\mu_2)(\nabla_3\mu_2)] \\ & + \exp(-\mu_3) \nabla_2[\exp(\mu_3)(\nabla_2\mu_3)]. \end{aligned} \quad (7.66)$$

7.2.2 Extrinsic Curvature and 4D Ricci Tensors

We also need various expressions for the extrinsic curvature (D.21)

$$(K^2)_{\hat{i}\hat{j}} = \frac{R^2}{4\alpha^2} \begin{pmatrix} (\nabla_2\omega)^2 + (\nabla_3\omega)^2 & 0 & 0 \\ 0 & (\nabla_2\omega)^2 & (\nabla_2\omega)(\nabla_3\omega) \\ 0 & (\nabla_2\omega)(\nabla_3\omega) & (\nabla_3\omega)^2 \end{pmatrix} \quad (7.67)$$

with its trace

$$\text{Tr}(K^2) = \frac{R^2}{2\alpha^2} [(\nabla_2\omega)^2 + (\nabla_3\omega)^2]. \quad (7.68)$$

For the calculation of the Lie derivative $(\mathcal{L}_\beta K)_{ij}$, we transform the extrinsic curvature to a coordinate basis and use the definition of the Lie derivative

$$(\mathcal{L}_\beta K)_{ij} = K_{ij,m}\beta^m + K_{im}\beta^m_{,j} + K_{mj}\beta^m_{,i}. \quad (7.69)$$

By transforming back to the orthonormal basis, we find for the Lie derivative of the extrinsic curvature

$$(\mathcal{L}_\beta K)_{\hat{i}\hat{j}} = \frac{R^2}{\alpha} \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\nabla_2\omega)^2 & (\nabla_2\omega)(\nabla_3\omega) \\ 0 & (\nabla_2\omega)(\nabla_3\omega) & (\nabla_3\omega)^2 \end{pmatrix}, \quad (7.70)$$

with its trace

$$\text{Tr}(\mathcal{L}_\beta K) = \frac{R^2}{\alpha} [(\nabla_2\omega)^2 + (\nabla_3\omega)^2]. \quad (7.71)$$

With these expressions, we can now derive the 4D Ricci tensors in orthonormal basis from the expressions (7.54)–(7.55)

$$R_{00} = \frac{1}{\alpha} \bar{\Delta}\alpha - \frac{R^2}{2\alpha^2} (\nabla\omega \cdot \nabla\omega) \quad (7.72)$$

$$R_{01} = \frac{1}{2} \text{Div} \left[\frac{R\nabla\omega}{\alpha} \right] \quad (7.73)$$

$$R_{02} = 0 \quad (7.74)$$

$$R_{03} = 0 \quad (7.75)$$

$$\begin{aligned} R_{11} = & -\frac{1}{R} \nabla_A (R\Psi_A) - \Psi_3 (\nabla_3\mu_2) - \Psi_2 (\nabla_2\mu_3) \\ & - \frac{R^2}{2\alpha^2} (\nabla\omega \cdot \nabla\omega) - \frac{1}{\alpha} (\nabla_A\Psi)(\nabla_A\alpha) \end{aligned} \quad (7.76)$$

$$\begin{aligned} R_{22} = & -\frac{1}{R} \nabla_2 (R\Psi_2) - \Psi_3 (\nabla_3\mu_2) - \Delta(\mu_2, \mu_3) \\ & + \frac{R^2}{2\alpha^2} (\nabla_2\omega)^2 - \frac{1}{\alpha} \nabla_2 (\nabla_2\alpha) - \frac{1}{\alpha} (\nabla_3\mu_2)(\nabla_3\alpha) \end{aligned} \quad (7.77)$$

$$\begin{aligned} R_{33} = & -\frac{1}{R}\nabla_3(R\Psi_3) - \Psi_2(\nabla_2\mu_3) - \Delta(\mu_2, \mu_3) \\ & + \frac{R^2}{2\alpha^2}(\nabla_3\omega)^2 - \frac{1}{\alpha}\nabla_3(\nabla_3\alpha) - \frac{1}{\alpha}(\nabla_2\mu_3)(\nabla_2\alpha) \end{aligned} \quad (7.78)$$

$$\begin{aligned} R_{23} = & -\frac{1}{R}\nabla_3(R\Psi_2) + \Psi_3(\nabla_2\mu_3) + \frac{R^2}{2\alpha^2}(\nabla_3\omega)(\nabla_2\omega) \\ & - \frac{1}{\alpha}(\nabla_3\mu_2)(\nabla_2\alpha) + \frac{1}{\alpha}\nabla_2(\nabla_3\alpha) \end{aligned} \quad (7.79)$$

$$R_{12} = 0 \quad (7.80)$$

$$R_{13} = 0. \quad (7.81)$$

7.2.3 3+1 Split of Einstein's Equations

As the 3D Ricci tensor has four nonvanishing components, we have found in total six equations for the five unknown functions α , ω , Ψ , μ_2 and μ_3 . The above decomposition of the Ricci tensor provides then the following equations:

- **Relativistic Poisson equation:** The equation for R_{00} results in an equation for the redshift factor, which we call the relativistic Poisson equation

$$\boxed{\bar{\Delta}\alpha - \frac{R^2}{2\alpha}(\nabla\omega \cdot \nabla\omega) = 8\alpha\pi G(T_{00} + T/2).} \quad (7.82)$$

- **Angular momentum equation:** the constraint equation for the extrinsic curvature determines the frame-dragging potential

$$\boxed{\text{Div}\left[\frac{R\nabla\omega}{\alpha}\right] = 16\pi GT_{0\phi}.} \quad (7.83)$$

- **Radius function:** The component R_{11} determines the cylindrical radius function

$$\begin{aligned} \frac{1}{R}\nabla_A[R\Psi_A] + \Psi_3\nabla_3\mu_2 + \Psi_2\nabla_2\mu_3 + \frac{R}{2\alpha^2}(\nabla\omega \cdot \nabla\omega) + \frac{1}{\alpha}(\nabla_A\Psi)(\nabla_A\alpha) \\ = -8\pi G[T_{11} - T/2]. \end{aligned} \quad (7.84)$$

This equation can also be written in closed form as (see problem 7.1)

$$\boxed{\frac{1}{\alpha}\text{Div}[\alpha\nabla\Psi] + \frac{R}{2\alpha^2}(\nabla\omega \cdot \nabla\omega) = -8\pi G[T_{11} - T/2].} \quad (7.85)$$

- **Meridional curvature:** the remaining two equations determine the curvature of the meridional plane, $\exp(\mu_2)$ and $\exp(\mu_3)$. With this we calculate the combination $R_{00} + R_{22} + R_{33} - R_{11}$

$$\begin{aligned} 2\Delta(\mu_2, \mu_3) - \frac{R^2}{2\alpha^2}(\nabla\omega \cdot \nabla\omega) - \frac{2}{\alpha}(\nabla_A\Psi)(\nabla_A\alpha) \\ = -8\pi G(T_{00} - T_{11} + T_{22} + T_{33}). \end{aligned} \quad (7.86)$$

In the isotropic gauge, $\mu_2 = \mu = \mu_3$, this is a second-order elliptic equation for μ alone.

- **Constraint equation:** Since there is no anisotropic source of matter, $T_{23} = 0$, $R_{23} = 0$ has to be satisfied identically.

7.3 Stellar Structure Equations in Isotropic Gauge

7.3.1 The Isotropic Gauge

Already Carter [106] has shown that the equations considerably simplify in the isotropic gauge (see also Bardeen [56])

$$\exp(2\mu_2) = \exp(2\mu_3) = \exp(2\mu). \quad (7.87)$$

One can discuss these either in cylindrical coordinates $d\varrho = dx^2$, $dz = dx^3$, or in pseudospherical coordinates $dr = dx^2$ and $r d\theta = dx^3$.

It is convenient to write the cylindrical radius as

$$\exp(\psi) = r \sin \theta B \exp(-v), \quad (7.88)$$

with B as a function of r and θ only. Originally, the stellar structure equations have been derived for a metric of the form

$$\begin{aligned} ds^2 = & -\exp(2v) c^2 dt^2 + B^2 r^2 \sin^2 \theta \exp(-2v) (d\phi - \omega dt)^2 \\ & + \exp(2\mu) (dr^2 + r^2 d\theta^2). \end{aligned} \quad (7.89)$$

There are two main effects that distinguish a rotating relativistic star from its nonrotating counterpart: The shape of the star is flattened by centrifugal forces (an effect that first appears at second order in the rotation rate), and the local inertial frames are dragged by the rotation of the source of the gravitational field. While the former effect is also present in the Newtonian limit, the latter is a purely relativistic effect. The study of the dragging of inertial frames in the spacetime of a rotating star is assisted by the introduction of the local Zero Angular-Momentum Observers (ZAMO). These are observers whose worldlines are normal to the $t = \text{const}$ hypersurfaces, and they are also called Eulerian observers. Then, the metric function ω is the angular velocity of the local ZAMO with respect to an observer at rest at infinity. Also, $\exp(-v)$ is the time dilation factor between the proper time of the local ZAMO and coordinate time t (which is the proper time at infinity) along a radial coordinate line. The metric function $\exp(\psi)$ has a geometrical meaning: it is the proper circumferential radius of a circle around the axis of symmetry. In the nonrotating limit, the metric (7.89) reduces to the metric of a nonrotating relativistic star in isotropic coordinates.

In rapidly rotating models, similar to black holes, an **ergosphere** can appear, where $g_{tt} > 0$. In this region, the rotational frame-dragging is strong enough to prohibit counter-rotating time-like or null geodesics to exist, and particles can have negative energy with respect to a stationary observer at infinity. Radiation fields (scalar, electromagnetic, or gravitational) can become unstable in the ergosphere.

The Rotating Matter

The four-velocity of matter is, $u^{(0)} = W$, and $u^{(1)} = W V$ in the locally nonrotating system, $\alpha = \exp v$,

$$u^{(0)} = \frac{W}{\alpha}, \quad u^{(1)} = \Omega u^{(0)}, \quad u^{(2)} = 0 = u^{(3)} \quad (7.90)$$

with the velocity V (in units of c) with respect to the Bardeen observer

$$V \equiv \frac{\varrho B(\Omega - \omega)}{\alpha^2} \quad (7.91)$$

and

$$W \equiv \frac{1}{\sqrt{1 - V^2}}. \quad (7.92)$$

The matter which is the source of the gravitational field has a total energy density ϵ and pressure P , and is given by the energy-momentum tensor

$$T^{\mu\nu} = (\epsilon + P)u^\mu u^\nu + Pg^{\mu\nu}. \quad (7.93)$$

In the locally nonrotating frame of reference, the energy-momentum tensor is given by

$$T^{(ab)} = \Theta_\mu^a \Theta_\nu^b T^{\mu\nu} \quad (7.94)$$

$$T^{(00)} = W^2(\epsilon + PV^2) \quad (7.95)$$

$$T^{(11)} = W^2(P + \epsilon V^2) \quad (7.96)$$

$$T^{(01)} = W^2(\epsilon + P)V \quad (7.97)$$

$$T^{(22)} = T^{(33)} = P \quad (7.98)$$

with its trace $T = -\epsilon + 3P$.

7.3.2 Structure Equations for Rotating Stars

Einstein's equations now follow from the Ricci tensor in the locally nonrotating frame, as derived in the previous section (they have been derived in this form for the first time by Butterworth and Ipser [94]). The equation for the redshift factor $\alpha = \exp(v)$ follows from R_{00} , $\varrho \equiv r \sin \theta$, equation (7.82),

$$\begin{aligned} \nabla_E \cdot [B \nabla_E v] - \frac{1}{2} \varrho^2 B^3 \exp(-4v) \nabla_E \omega \cdot \nabla_E \omega \\ = 4\pi G \exp(2\mu) [W^2(\epsilon + P)(1 + V^2) + 2P]. \end{aligned} \quad (7.99)$$

The differential operators ∇_E and the divergence $\nabla_E \cdot$ are taken in flat Euclidean three-space. With this meaning, the equations can be written down in any coordinate

system x^2 and x^3 in the isotropic gauge. This is essentially the Poisson equation for the generalized Newtonian potential v including the fact that self-gravity from the frame-dragging potential is also a source for the redshift factor. A second equation can be written down for the frame-dragging potential $\omega(r, \theta)$ from the expression (7.83)

$$\nabla_E \cdot [\varrho^2 B^3 \exp(-4v) \nabla_E \omega] = -16\pi G \varrho B^2 \exp(2\mu - 2v) W^2 (\epsilon + P) V. \quad (7.100)$$

This shows explicitly that the matter current generated by rotation is the source for the frame-dragging potential. The equation for B is also a simple divergence equation and follows from (7.85)

$$\nabla_E \cdot [\varrho \nabla_E B] = 16\pi G \varrho B \exp(2\mu) P. \quad (7.101)$$

These three equations determine the redshift factor, the frame-dragging potential and the radius function.

The cylindrical radius function is therefore given by the pressure. What is the source of the fourth function $\exp(\mu)$? In the above equations, this function only appears in the determinant of the metric. Essentially, we have two additional equations at our disposal which have not yet been used, R_{22} and R_{23} . As in the static case, one equation hides the hydrostatic equilibrium, and the other one is responsible for the function μ . It turns out that this equation for μ only depends on the first three metric functions v , B and ω by means of first-order derivatives. It results in a quite lengthy expression (see Butterworth and Ipser [94]). Thus, three of the four gravitational field equations are elliptic, while the fourth equation is a first-order partial differential equation, relating only metric functions. Alternatively, one could use equation (7.86) as an elliptic second-order equation for μ

$$\begin{aligned} \nabla_E^2 \mu - \frac{R^2}{2\alpha^2} (\nabla_E \omega \cdot \nabla_E \omega) - \frac{2}{\alpha} (\nabla_E \Psi) (\nabla_E \alpha) \\ = -8\pi G (\epsilon + P) \exp(2\mu). \end{aligned} \quad (7.102)$$

The exterior metric of a rapidly rotating neutron star considerably differs from the Kerr metric. The two metrics only agree to lowest order in the rotational velocity. At higher order, the multipole moments of the gravitational field created by a rapidly rotating compact star are different from the multipole moments of the Kerr field. There have been many attempts in the past to find analytic solutions to the Einstein equations in the stationary, axisymmetric case, that could describe a rapidly rotating neutron star. An interesting solution has been found by Manko et al. (see Sect. 7.6). This solution reduces to a three-parameter solution, involving the mass, specific angular momentum, and a parameter that depends on the quadrupole moment of the source. Although this solution explicitly depends only on the quadrupole moment, it approximates the gravitational field of a rapidly rotating star with higher nonzero multipole moments. It will be interesting to determine whether this analytic quadrupole solution approximates the exterior field of a rapidly rotating star more accurately than the quadrupole, slow rotation approximation.

7.3.3 Mechanical Equilibrium and Effective Potential

In this section we derive the equations of motion for plasma in a general axisymmetric spacetime. Due to axisymmetry, we have two global conservation laws, the conservation of baryon number and the conservation of angular momentum. The hydrostatic equilibrium follows then as a special case of the poloidal equations of motion.

As discussed in Sect. 3.1, the motion of a plasma in the metric (7.89) is defined by the following conserved variables

$$D = \varrho_0 W \quad (7.103)$$

$$S_A = \varrho_0 h W^2 v_A \quad (7.104)$$

$$S_\phi = \varrho_0 h W^2 v_\phi \quad (7.105)$$

$$\tau = E - D, \quad (7.106)$$

where we used the following primary variables

$$v^A = \frac{u^A}{W} + \frac{\beta^A}{\alpha} \quad (7.107)$$

$$v_\phi = \frac{u_\phi}{W} \quad (7.108)$$

$$\Omega = u^\phi/u^t \quad (7.109)$$

$$W = \alpha u^t = 1/\sqrt{1 - v^A v_A - v^\phi v_\phi}. \quad (7.110)$$

S_ϕ is the angular momentum density of the plasma, while u_ϕ is the specific angular momentum with its relation

$$u_\phi = \frac{S_\phi}{\varrho_0 h W}. \quad (7.111)$$

The quantities S_A are related to meridional motion v^A , $A = r, \theta$.

The equation of hydrostationary equilibrium follows from the projection of the conservation of the stress–energy tensor normal to the four-velocity, or from the two equations for the poloidal momentum density S_A in (3.32)

$$\frac{1}{\sqrt{-g}} \partial_B \left[\sqrt{-g} \mathbf{F}^B[S_A] \right] = \delta[S_A], \quad (7.112)$$

where $\mathbf{F}^B[S_A] = P \delta_A^B$ is the only nonvanishing flux when $v^A = 0$. This leads to two poloidal equations

$$P_{,A} + \frac{P}{\sqrt{-g}} \partial_A \left[\sqrt{-g} \right] = T^{BC} \partial_B (g_{CA}) - T^{\mu\nu} \Gamma_{\mu\nu}^B g_{BA}. \quad (7.113)$$

In terms of Ω , this mechanical equilibrium can be written in a compact form (for the derivation, see problem 7.4), using $u^t = W/\alpha$,

$$\boxed{\nabla_A P - (\varrho + P) [\nabla_A \ln u^t - u^t u_\phi \nabla_A \Omega] = 0.} \quad (7.114)$$

∇ is the nabla operator in the two-dimensional meridional plane, and we assume some EoS $P = P(\varrho)$ given. This can be written with the specific angular momentum $l = -u_\phi/u_t$

$$\nabla P - (\varrho + P) [\nabla \ln u^t + u^t u_t l \nabla \Omega] = 0, \quad (7.115)$$

and due to the normalization $u^t u_t + u^\phi u_\phi = -1$ we have

$$u^t u_t (1 - l\Omega) = -1. \quad (7.116)$$

Therefore the hydrostatic equilibrium follows as

$$\frac{\nabla P}{\varrho + P} = \nabla \ln u^t - \frac{l}{1 - l\Omega} \nabla \Omega, \quad (7.117)$$

or explicitly

$$\boxed{\frac{\nabla P}{\varrho + P} = -\nabla \ln \alpha + \nabla \ln W - \frac{l}{1 - l\Omega} \nabla \Omega.} \quad (7.118)$$

The first term is the gravitational force as measured in the ZAMO, the second one comes from the motion, and the third term is due to differential rotation.

For barotropic equations of state, $P = P(\varrho_0)$, this can easily be integrated to yield

$$\int_0^P \frac{dP}{\varrho + P} = \ln u^t - \ln u_0^t - F(\Omega) \quad (7.119)$$

with

$$F(\Omega) \equiv \int_{\Omega_c}^{\Omega} \frac{l(\Omega')}{1 - \Omega' l(\Omega')} d\Omega'. \quad (7.120)$$

Ω_c is the value of Ω on the symmetry axis. This function $F(\Omega)$ vanishes for rigid rotation. Here, we use the relativistic enthalpy $\mu \equiv (\varrho + P)/\varrho_0$ and the specific angular momentum $l = l(\Omega)$ given by the inversion of $\Omega(l)$

$$l(\Omega) = \frac{R^2(\Omega - \omega)}{\alpha^2 + R^2\omega(\Omega - \omega)/c^2}. \quad (7.121)$$

This is the classical expression for the specific angular momentum in the limit $\alpha \rightarrow 1$ and far away from the central source, $l \rightarrow R^2\Omega$.

This means that the pressure gradient is given by the **effective potential** Ψ_{eff} defined as

$$\Psi_{\text{eff}} \equiv \ln u^t + F(\Omega) = \ln(W/\alpha) + F(\Omega), \quad (7.122)$$

with

$$\boxed{\frac{1}{\varrho + P} \nabla P = \nabla \Psi_{\text{eff}}.} \quad (7.123)$$

This generalizes the expression of the effective potential in Newtonian physics

$$\Psi_{\text{eff}} = \Phi_N + \int_0^R R' \Omega^2(R') dR' \quad (7.124)$$

with the associated hydrostatic equilibrium

$$\frac{1}{\varrho_0} \nabla P = -\nabla \Psi_{\text{eff}}. \quad (7.125)$$

Here, we have assumed that the rotation is constant on cylinders.

7.3.4 Stellar Parameters

Similar to static neutron stars, certain integrals over the matter distribution are observable quantities. For a rapidly rotating neutron star, we can compute numerically the physical quantities shown in Table 7.1, once the above structure equations have been solved. For this purpose, we assume the metric is given in the following form

$$ds^2 = -\exp(2\Phi) c^2 dt^2 + \exp(2\beta) r^2 \sin^2 \theta (d\phi - \omega dt)^2 + \exp(2\mu) (dr^2 + r^2 d\theta^2). \quad (7.126)$$

Rest Mass and Proper Mass of a Star

First of all we can define the rest mass of the star (baryon number) in terms of the rest-mass density ϱ_0

$$M_0 = \int \varrho_0 u^\mu d\Sigma_\mu = \int \varrho_0 u^t \sqrt{-g} d^3x = \int \varrho_0 \frac{W}{\alpha} \sqrt{-g} d^3x = \int \varrho_0 W \sqrt{\gamma} d^3x, \quad (7.127)$$

or in coordinates

$$M_0 = 2\pi \int \int \varrho_0 \frac{e^{2\mu+\beta}}{\sqrt{1-v^2}} r^2 \sin \theta dr d\theta. \quad (7.128)$$

The volume element $d\Sigma_\mu = k_\mu \sqrt{-g} d^3x$ can be expressed as a normal volume element. The rest mass itself does not appear as a source for gravitational fields. Similarly, we find for the proper mass

$$M_p = 2\pi \int \int \varepsilon \frac{e^{2\mu+\beta}}{\sqrt{1-v^2}} r^2 \sin \theta dr d\theta. \quad (7.129)$$

Table 7.1. Parameters of rotating stars

Parameter	Physical meaning
ε_c	Central energy density
r_p/r_e	Ratio of polar to equatorial radii
Ω	Angular velocity of the star
M_0	Baryon mass
M_p	Proper mass
M	Gravitational mass
R_{circ}	Equatorial circumferential radius
r_e	Equatorial coordinate radius
J	Total angular momentum
I	Moment of inertia about the rotation axis
T	Rotational energy
W_G	Gravitational energy
v_e	Velocity of comoving observer at the equator relative to the locally nonrotating observer
Z_p	Polar redshift
Z_c	Central redshift
Z_{eq}^b	Equatorial redshift in the backward direction
Z_{eq}^f	Equatorial redshift in the forward direction
e	Intrinsic eccentricity of the star's surface
$GRV2$	Two dimensional virial identity
$GRV3$	Three dimensional virial identity

Mass and Angular Momentum for Axisymmetric Systems

We have seen that in suitable asymptotically Minkowskian coordinates, such that the stationary Killing vector has the form $k = \partial_t$, the metric components are always given by

$$g_{00} = -1 + (2GM/r) + O(1/r^2) \quad (7.130)$$

$$g_{0\phi} = (2GJ/r^3) + O(1/r^4) \quad (7.131)$$

$$g_{ij} = [1 + (2GM/r)] \delta_{ij} + O(1/r^2). \quad (7.132)$$

The meaning of the constants M and J is a consequence of the Komar integrals [234]. Let V be a volume of spacetime on a space-like hypersurface Σ with boundary ∂V . To every Killing vector field ξ we can associate the Komar integral

$$Q_\xi(V) = \frac{q}{16\pi G} \oint_{\partial V} dS_{\mu\nu} D^\mu \xi^\nu \quad (7.133)$$

for some constant q . Using Gauss' law, we find

$$Q_\xi(V) = \frac{q}{8\pi G} \int_V d\Sigma_\mu D_\nu (D^\mu \xi^\nu). \quad (7.134)$$

Since Killing fields satisfy the identity

$$D_\nu D_\mu \xi^\nu = R_{\mu\nu} \xi^\nu, \quad (7.135)$$

we obtain over Einstein's equations

$$Q_\xi(V) = \frac{q}{8\pi G} \int_V d\Sigma_\mu R_\nu^\mu \xi^\nu = \int_V dS_\mu J_{(\xi)}^\mu, \quad (7.136)$$

where the current $J_{(\xi)}^\mu$ is defined as follows

$$J_{(\xi)}^\mu = q \left(T_\nu^\mu \xi^\nu - \frac{1}{2} T \xi^\mu \right). \quad (7.137)$$

One can easily prove that his current is conserved, i.e. $D_\mu J_\xi^\mu = 0$. Since this current is conserved, the charge $Q_\xi(V)$ is time-independent, provided J^μ vanishes on the boundary ∂V .

For $\xi = k, q = -2$ is fixed by comparison with the Schwarzschild metric, and

$$M = M(V) = -\frac{1}{8\pi G} \oint_{\partial V} dS_{\mu\nu} D^\mu k^\nu \quad (7.138)$$

is the total energy contained in the volume V . For $\xi = m$, where $m = \partial_\phi$, we obtain for $q = 1$ the angular momentum integral

$$J = J(V) = \frac{1}{16\pi G} \oint_{\partial V} dS_{\mu\nu} D^\mu m^\nu. \quad (7.139)$$

To check the coefficient, use Gauss' law to write $J = \int_V dS_\mu J_m^\mu$, where the angular momentum current is given by

$$J_{(m)}^\mu = T_\nu^\mu m^\nu - \frac{1}{2} T m^\mu. \quad (7.140)$$

Since $dS_\mu m^\mu = 0$, we have

$$J(V) = \int_V dV T_\nu^0 m^\nu = \int_V dV (T_2^0 x^1 - T_1^0 x^2) \quad (7.141)$$

in Cartesian coordinates, where

$$m = x^1 \partial_2 - x^2 \partial_1. \quad (7.142)$$

For weak gravitational sources, $g \simeq \eta$, we have

$$J(V) \simeq \epsilon_{3jk} \int_V d^3x x^j T^{k0}. \quad (7.143)$$

Therefore, the total angular momentum of an asymptotically flat spacetime is found by taking ∂V to be a two-sphere at spatial infinity and calculating the integral in (7.139).

According to these considerations, the two quantities M and J as defined above by the asymptotic expansion can be given a coordinate-independent form

$$8\pi GM = - \lim_{\infty} \oint (\nabla^\alpha k^\beta) dS_{\alpha\beta} \quad (7.144)$$

$$16\pi GJ = \lim_{\infty} \oint (\nabla^\alpha m^\beta) dS_{\alpha\beta} \quad (7.145)$$

where $m = \partial_\phi$ is the Killing vector of axisymmetry and where the integral is taken over a space-like two-surface S with $dS = (e_\phi \wedge e_0) dS$ for the two mutually orthogonal space-like and time-like normals e_ϕ and e_0 . The limit is taken over a two-sphere at arbitrarily large asymptotic distances. One can then use the identities

$$(\nabla_\alpha \nabla^\alpha) k^\beta = R^\beta_\varrho k^\varrho, \quad (\nabla_\alpha \nabla^\alpha) m^\beta = R^\beta_\varrho m^\varrho \quad (7.146)$$

which hold in consequence of the Killing equations

$$\nabla_{(\alpha} k_{\beta)} = 0 = \nabla_{(\alpha} m_{\beta)} \quad (7.147)$$

to convert these expressions by means of Stokes' theorem to the form

$$M = - \frac{1}{4\pi G} \int_{\Sigma} R^\alpha_\beta k^\beta d\Sigma_\alpha \quad (7.148)$$

$$J = \frac{1}{8\pi G} \int_{\Sigma} R^\alpha_\beta m^\beta d\Sigma_\alpha. \quad (7.149)$$

Σ is a space-like hypersurface extending from space-like infinity to the origin of the star (or some bounding horizon in the case of black holes). Inserting Einstein's equations

$$R^\alpha_\beta = 8\pi G \left(T^\alpha_\beta - \frac{1}{2} \text{Tr}(T) g^\alpha_\beta \right) \quad (7.150)$$

we see that the two constants are given by

$$M = - \int_{\Sigma} (2T^\alpha_\beta k^\beta - \text{Tr}(T) k^\alpha) d\Sigma_\alpha \quad (7.151)$$

$$J = \int_{\Sigma} T^\alpha_\beta m^\beta d\Sigma_\alpha. \quad (7.152)$$

The gravitational source is the total mass of isolated systems which can be expressed as an integral over a space-like slice

$$M = - \int (2T^\mu_v - \delta^\mu_v T) k^v d\Sigma_\mu = \int_V (2T^t_t + T) \sqrt{-g} d^3x, \quad (7.153)$$

or explicitly

$$M = 2\pi \int \int \left[e^{2\mu+\beta} \left\{ \frac{(\varepsilon + p)(1 + v^2)}{1 - v^2} + 2p \right\} + 2r \sin \theta \omega e^\beta \frac{(\varepsilon + p)v}{1 - v^2} \right] r^2 \sin \theta dr d\theta. \quad (7.154)$$

This mass appears as parameter in the redshift factor with its asymptotic expansion

$$\alpha(r) = 1 - \frac{GM}{c^2 r} + O(1/r^3). \quad (7.155)$$

Similarly, the total angular momentum of the star can be expressed as an integral over the corresponding source term

$$J = \int T_\beta^\alpha m^\beta d\Sigma_\alpha = \int T_\phi^t \sqrt{-g} d^3x = \int j dM_0, \quad (7.156)$$

or

$$J = 2\pi \int \int e^{2\mu+2\beta} \frac{(\varepsilon + p)v}{1 - v^2} r^3 \sin^2 \theta dr d\theta. \quad (7.157)$$

This quantity appears as the source for the frame-dragging potential in the asymptotic region

$$\omega = \frac{2GJ}{c^3 r^3} + O(1/r^4). \quad (7.158)$$

The angular momentum determines the moment of inertia by means of $I = J/\Omega$. From this expression, we also get the total kinetic energy

$$T = \frac{1}{2} \int \int \Omega dJ = 2\pi \int \int e^{2\mu+2\beta} \frac{(\varepsilon + p)v}{1 - v^2} \Omega r^3 \sin^2 \theta dr d\theta, \quad (7.159)$$

as well as the gravitational energy

$$W_G = M_p c^2 + T - Mc^2. \quad (7.160)$$

Redshifts and Ellipsoidal Structure

The various redshifts from the surface can easily be calculated by means of

$$Z_p = \exp(-\Phi_p) - 1, \quad (7.161)$$

$$Z_{\text{eq}}^f = \left(\frac{1 - v_e}{1 + v_e} \right)^{1/2} \frac{\exp(-\Phi_e)}{1 + r_e \exp[(\Phi_e - \beta_e)/2]\omega_e} - 1, \quad (7.162)$$

$$Z_{\text{eq}}^b = \left(\frac{1 + v_e}{1 - v_e} \right)^{1/2} \frac{\exp(-\Phi_e)}{1 - r_e \exp[(\Phi_e - \beta_e)/2] \omega_e} - 1, \quad (7.163)$$

where subscripts p and e denote values at the pole and the equatorial surface, respectively.

The eccentricity of the meridional cross-section is defined by the following procedure (Friedman et al. [161]). If the surface of the star is defined by

$$r = r_s(\theta), \quad (7.164)$$

the metric of the stellar surface can be expressed as

$$d\sigma_*^2 = e^{2\beta} r^2 \sin^2 \theta d\varphi^2 + e^{2\mu} \left[\left(\frac{dr_s}{d\theta} \right)^2 + r_s^2(\theta) \right] d\theta^2. \quad (7.165)$$

If we embed this surface in the flat three-dimensional space, it is expressed as

$$R = R_s(z), \quad (7.166)$$

in cylindrical coordinates (R, φ, z) . The two-metric of this surface is

$$d\sigma_*^2 = \left[\left(\frac{dR_s}{dz} \right)^2 + 1 \right] dz^2 + R_s^2 d\varphi^2. \quad (7.167)$$

Comparing these two equations, we have the following relations, if they express the same surface geometry

$$R_s(\theta) = e^\beta r \sin \theta, \quad (7.168)$$

and

$$z_s(\theta) = \int_\theta^{\pi/2} d\theta \left\{ e^{2\mu} \left[\left(\frac{dr_s}{d\theta} \right)^2 + r_s(\theta)^2 \right] - \left(\frac{dR_s}{d\theta} \right)^2 \right\}^{1/2}. \quad (7.169)$$

Using these quantities, the eccentricity e is defined as

$$e \equiv \sqrt{1 - \left(\frac{z_s(\theta = 0)}{R_s(\theta = \pi/2)} \right)^2}. \quad (7.170)$$

Binding Energy

When we introduce a quantity h in the star as

$$\frac{\mu}{U^0} = 1 - h \quad (7.171)$$

then we get for the binding energy of the object

$$E_B \equiv M_0 - M = \int h \, dM_0 - 2 \int j\Omega \, dM_0 - 2 \int \frac{P}{\varrho_0 u^t} \, dM_0. \quad (7.172)$$

The surfaces of constant h must coincide with the surfaces of constant Ω . This can be expressed differentially: if the rest mass changes by the amount δdM_0 and the angular momentum by $\delta J = j \, dM_0$, the change in the binding energy is

$$\delta E_B = \int h \, \delta dM_0 - \int \Omega \, \delta dM_0. \quad (7.173)$$

Similar to a static star, an equilibrium model is uniquely specified by the isentropic equation of state, $P = P(\epsilon)$, the value of the total rest mass M_0 and either by the angular momentum per unit rest mass or the angular velocity Ω of the stellar surface [55].

7.4 The Slow-Rotation Approximation

In the slow-rotation limit, we can write the line element as

$$ds^2 = -\exp(2\Phi) dt^2 + \left(1 - \frac{2M(r)}{r}\right)^{-1} dr^2 - 2\omega r^2 \sin^2 \theta \, dt \, d\phi + r^2 d\Omega^2.$$

(7.174)

As in the nonrotating case, $M(r)$ and Φ satisfy the TOV equations

$$\frac{dM}{dr} = 4\pi r^2 \varrho \quad (7.175)$$

$$\frac{d\Phi}{dr} = \frac{M(r) + 4\pi r^3 P}{1 - 2M(r)/r}. \quad (7.176)$$

The frame-dragging potential satisfies the equation

$$\frac{1}{r^3} \frac{d}{dr} \left(r^4 \iota \frac{d\bar{\omega}}{dr} \right) + 4 \frac{d\iota}{dr} \bar{\omega} = 0, \quad (7.177)$$

where $\bar{\omega} = \Omega - \omega$ is the rotational frequency with respect to nonrotating frames, and $\iota(r)$ is defined as

$$\iota = \sqrt{1 - 2M(r)/r} \exp(-\Phi). \quad (7.178)$$

The angular momentum J of the star follows from

$$J = \int_M T_v^\mu m^\nu d^3 \Sigma_\mu \quad (7.179)$$

with $m = \partial_\phi$ as the axial Killing field and the four-velocity $u = u^t(1, 0, 0, \Omega)$. $\Omega = u^\phi/u^t$ is the angular velocity of the star. From this expression we obtain

$$J = \int_M (\varrho + P)(u^t)^2 (g_{t\phi} + \Omega g_{\phi\phi}) \sqrt{-g} d^3x \quad (7.180)$$

or

$$J = \int_M (\varrho + P)(u^t)^2 g_{\phi\phi}(\Omega - \omega) \sqrt{-g} d^3x. \quad (7.181)$$

This allows us to define the moment of inertia by means of $I = J/\Omega$. In leading order in the expansion for Ω we obtain

$$I = \frac{8\pi}{3} \int_0^R r^4 \frac{(\varrho + P) \exp(-\Phi)}{\sqrt{1 - 2M(r)/r}} \frac{\bar{\omega}}{\Omega} dr. \quad (7.182)$$

The Mass-Quadrupole Moment

Rotation also leads to a quadrupolar deformation of the surfaces of constant density. We have therefore also to perturb the redshift factor and the volume element. Hartle and Thorne [195] have derived an exact form of this metric with accuracy up to second-order terms in the specific angular momentum $j = c^3 J/G^2 M^2$, and first order in the dimensionless mass-quadrupole moment $q = Q/M^3$, where M denotes the mass of the star, J its angular momentum and Q its quadrupole mass moment. Since $J = I\Omega$ and $I \simeq 0.25MR_*^2$, the dimensionless quantity j can be estimated in terms of the Keplerian rotation at the equatorial radius, Ω_K ,

$$j \simeq 0.25 \sqrt{R_*/M} (\Omega/\Omega_K) < 0.2 \quad (7.183)$$

for observed neutron stars with periods $P > 1.5$ ms. In distinction to Schwarzschild, this solution now has three parameters. We use geometrical units with $c = 1 = G$. In these units, the mass is given in terms of the gravitational radius $M = GM_*/c^2$, and the Hartle–Thorne (HT) metric

$$ds_{\text{HT}}^2 = -\alpha^2 dt^2 + \gamma_{\phi\phi}(d\phi - \omega dt)^2 + \gamma_{rr} dr^2 + \gamma_{\theta\theta} d\theta^2, \quad (7.184)$$

where

$$\alpha^2 = \Delta R, \quad \omega = \frac{2J}{r^3} \quad (7.185)$$

$$\gamma_{rr} = \frac{S}{\Delta}, \quad \gamma_{\theta\theta} = r^2 B \quad (7.186)$$

$$\gamma_{\phi\phi} = r^2 \sin^2 \theta B \quad (7.187)$$

$$g_{tt} = -[\alpha^2 - \gamma_{\phi\phi}\omega^2] = -\left[\Delta R - \frac{4J^2}{r^4} B \sin^2 \theta\right] \quad (7.188)$$

$$g_{t\phi} = -\omega\gamma_{\phi\phi} = -\frac{2J}{r} B \sin^2 \theta. \quad (7.189)$$

The functions are given as follows

$$\Delta = 1 - \frac{2M}{r} + \frac{2J^2}{r^4} \quad (7.190)$$

$$R = 1 + 2 \left[\frac{J^2}{Mr^3} \left(1 + \frac{M}{r} \right) + \frac{5}{8} \frac{Q - J^2/M}{M^3} Q_2^2(r/M - 1) \right] P_2 \quad (7.191)$$

$$S = 1 - 2 \left[\frac{J^2}{Mr^3} \left(1 - \frac{5M}{r} \right) + \frac{5}{8} \frac{Q - J^2/M}{M^3} Q_2^2(r/M - 1) \right] P_2 \quad (7.192)$$

$$B = 1 + 2 \left[-\frac{J^2}{Mr^3} \left(1 + \frac{2M}{r} \right) + \frac{5}{8} \frac{Q - J^2/M}{M^3} \times \left\{ \frac{2M}{\sqrt{r^2(1 - 2M/r)}} Q_2^1(r/M - 1) - Q_2^2(r/M - 1) \right\} \right] P_2. \quad (7.193)$$

M is the mass, J the angular momentum and Q the mass quadrupole moment, $P_2 \equiv P_2(\cos \theta) = (3 \cos^2 \theta - 1)/2$ is the second Legendre polynomial, and Q_n^m the associated Legendre polynomials, defined as

$$Q_2^1(z) = \sqrt{1-z^2} \left[\frac{3z^2-2}{z^2-1} - \frac{3}{2}z \ln \left(\frac{z+1}{z-1} \right) \right] \quad (7.194)$$

$$Q_2^2(z) = \frac{3}{2}(z^2-1) \ln \left(\frac{z+1}{z-1} \right) - \frac{3z^3-5z}{z^2-1}. \quad (7.195)$$

This gives the following explicit expressions which can numerically be used

$$Q_2^1(r/M - 1) = \frac{r}{M} \sqrt{1 - \frac{2M}{r}} \times \left[\frac{3(r/M)^2(1 - 2M/r) + 1}{(r/M)^2(1 - 2M/r)} + \frac{3}{2} \frac{r}{M} \left(1 - \frac{r}{M} \right) \ln \left(1 - \frac{2M}{r} \right) \right] \quad (7.196)$$

$$Q_2^2(r/M - 1) = -\frac{3}{2} \left(\frac{r}{M} \right)^2 \left(1 - \frac{2M}{r} \right) \ln \left(1 - \frac{2M}{r} \right) - \frac{M/r(1 - M/r)[3(r/M)^2(1 - 2M/r) - 2]}{1 - 2M/r}. \quad (7.197)$$

The Hartle–Thorne metric is sometimes used in a somewhat different notation

$$ds_{\text{HT}}^2 = -\exp \nu [1 + 2\Phi(r, \theta)] dt^2 - \exp \lambda [1 + 2M/(r - 2M)] dr^2 + r^2 [1 + 2\Psi(r, \theta)][d\theta^2 + \sin^2 \theta (d\phi - \omega dt)^2] \quad (7.198)$$

can be expressed in spherical coordinates (t, r, θ, ϕ)

$$g_{tt} = -(1 - 2M/r)[1 + j^2 F_1 - qF_2] \quad (7.199)$$

$$g_{rr} = (1 - 2M/r)^{-1}[1 + j^2 G_1 + qF_2] \quad (7.200)$$

$$g_{\theta\theta} = r^2[1 + j^2 H_1 - qH_2] \quad (7.201)$$

$$g_{\phi\phi} = g_{\theta\theta} \sin^2 \theta \quad (7.202)$$

$$g_{t\phi} = 2(jM^2/r) \sin^2 \theta. \quad (7.203)$$

The dimensionless angular momentum j and quadrupole moment q are defined by

$$j = J/M^2, \quad q = Q/M^3. \quad (7.204)$$

Here, we use the following functions, $\mu = \cos \theta$,

$$\begin{aligned} F_1 &= [8Mr^4(r - 2M)]^{-1} \times \\ &[\mu^2(48M^6 - 8M^5r - 24M^4r^2 - 30M^3r^3 - 60M^2r^4 + 135Mr^5 - 45r^6) \\ &+ (r - M)(16M^5 + 8M^4r - 10M^2r^3 - 30Mr^4 + 15r^5)] + A_1(r) \end{aligned} \quad (7.205)$$

$$\begin{aligned} F_2 &= [8Mr(r - 2M)]^{-1}(5(3\mu^2 - 1)(r - M)(2M^2 + 6Mr - 3r^2)) \\ &- A_1(r) \end{aligned} \quad (7.206)$$

$$G_1 = [8Mr^4(r - 2M)]^{-1}((L - 72M^5r) - 3\mu^2(L - 56M^5r)) - A_1(r) \quad (7.207)$$

$$L = 80M^6 + 8M^4r^2 + 10M^3r^3 + 20M^2r^4 - 45Mr^5 + 15r^6 \quad (7.208)$$

$$A_1 = \frac{15r(r - 2M)(1 - 3\mu^2)}{16M^2} \ln\left(\frac{r}{r - 2M}\right) \quad (7.209)$$

$$\begin{aligned} H_1 &= (8Mr^4)^{-1}(1 - 3\mu^2)(16M^5 + 8M^4r - 10M^2r^3 + 15Mr^4 + 15r^5) \\ &+ A_2(r) \end{aligned} \quad (7.210)$$

$$H_2 = (8Mr)^{-1}(5(1 - 3\mu^2)(2M^2 - 3Mr - 3r^2)) - A_2(r) \quad (7.211)$$

$$A_2 = \frac{15(r^2 - 2M^2)(3\mu^2 - 1)}{16M^2} \ln\left(\frac{r}{r - 2M}\right). \quad (7.212)$$

The Kerr metric in Boyer–Lyndquist coordinates could be obtained from the above metric after putting $a = jM$ and $q = j^2$, and making a coordinate transformation of the form

$$r_{BL} = r - a^2/(2r^3)((r + 2M)(r - M) + \mu^2(r - 2M)(r + 3M)) \quad (7.213)$$

$$\theta_{BL} = \theta - a^2/(2r^3)(3 + 2M) \cos \theta \sin \theta. \quad (7.214)$$

7.5 Numerical Integration of the Stellar Structure Equations

As we have seen, the stellar structure is essentially determined by four elliptic partial differential equations and the equation of hydrostatic equilibrium which yields us the density structure in the interior. First models have been calculated by Wilson (1972) and Bonazzola and Schneider (1974). Modern treatments essentially rely on the following methods (see also Stergioulas [376]):

- Slow rotation approximation discussed in the previous section: This method includes quadrupole corrections to spherically symmetric stars induced by Ω^2 . This method is definitely not appropriate for the treatment of rapid rotation in millisecond pulsars.
- Butterworth and Ipser [94] applied a linearization procedure starting with non-rotating configurations. This is also not suitable for rapid rotation.
- **KEH:** The most advanced method is based on a Green's function approximation (Komatsu, Eriguchi and Hachisu [235] (KEH)). The essential procedure is based on the fact that the three elliptic type field equations can be converted into integral equations using corresponding Green's functions. As we have seen, the fourth field equation is an ordinary first-order differential equation, which can simply be solved. Cook, Shapiro and Teukolsky [123, 124] have improved this KEH scheme by inventing a new radial coordinate

$$r = r_e \frac{s}{1-s}, \quad (7.215)$$

mapping in this way the entire radial range into a finite interval. For this reason, the region of integration has not to be truncated. This leads to a much faster convergence. This code is public domain software and is described in the PhD thesis of Stergioulas (SF code, [375]).

- Bonazzola and collaborators have developed a different method (spectral method): functions are expanded in trigonometric functions in r and θ so that fast Fourier transform methods can be applied [83, 85] (BGSM).

BGSM, Lorene

In the BGSM scheme, the field equations are derived in the 3+1 formulation with the identification

$$\exp v = N, \quad \omega = N^\phi \quad (7.216)$$

$$\exp 2\psi = A^4 B^2 r^2 \sin \theta \quad (7.217)$$

$$\exp \mu_2 = A^2/B, \quad \exp \mu_3 = rA^2/B. \quad (7.218)$$

All four chosen equations that describe the gravitational field are of elliptic type. This avoids the problem with the second-order radial derivative in the source term of the ODE used in BI and KEH. The equations are solved using a spectral method, i.e. all functions are expanded in terms of trigonometric functions in both the angular and radial directions and a Fast Fourier Transform (FFT) is used to obtain coefficients. Outside the star, a redefined radial variable is used, which maps infinity to a finite distance. In Salgado et al. [352, 353] the code is used to construct a large number of models based on various EoSs. The accuracy of the computed models is estimated using two general relativistic virial identities, valid for general asymptotically flat spacetimes. While the field equations used in the BI and KEH schemes assume a perfect fluid, isotropic stress-energy tensor, the BGSM formulation makes no

assumption about the isotropy of the energy–momentum tensor. Thus, the BGSM code can compute stars with a magnetic field, a solid crust, or a solid interior, and it can also be used to construct rotating boson stars.

Bonazzola et al. [85] have improved the BGSM spectral method by allowing for several domains of integration. One of the domain boundaries is chosen to coincide with the surface of the star and a regularization procedure is introduced for the divergent derivatives at the surface (that appear in the density field when stiff equations of state are used). This allows models to be computed that are nearly free of Gibbs phenomena at the surface. The same method is also suitable for constructing quasistationary models of binary neutron stars. The new method has been used in Gourgoulhon [181] for computing models of rapidly rotating strange stars and it has also been used in 3D computations of the onset of the viscosity-driven instability to bar-mode formation.

AKM

A new multidomain spectral method has been introduced in Ansorg et al. (AKM, [34, 35]). The method can use several domains inside the star, one for each possible phase transition. Surface-adapted coordinates are used and approximated by a two-dimensional Chebyshev expansion. Requiring transition conditions to be satisfied at the boundary of each domain, the field and fluid equations are solved as a free boundary value problem by a Newton–Raphson method, starting from an initial guess. The field equations are simplified by using a corotating reference frame. Applying this new method to the computation of rapidly rotating homogeneous relativistic stars, AKM achieve near machine accuracy, except for configurations at the mass-shedding limit. The code has been used in a systematic study of uniformly rotating homogeneous stars in general relativity (Schöbel and Ansorg [358]).

7.5.1 Comparison of Numerical Codes

The accuracy of the above numerical codes can be estimated, if one constructs exactly the same models with different codes and compares them directly. The first such comparison of rapidly rotating models constructed with the FIP and SF codes has been presented by Stergioulas and Friedman [377]. Rapidly rotating models constructed with several EoS’s agree to 0.1%–1.2% in the masses and radii and to better than 2% in any other quantity that was compared (angular velocity and momentum, central values of metric functions, etc.). This is a very satisfactory agreement, considering that the BI code was using relatively few grid points, due to limitations of computing power at the time of its implementation.

If one makes the extreme assumption of uniform density, the agreement is at the level of 10^{-2} . In the BGSM code this is due to the fact that the spectral expansion in terms of trigonometric functions cannot accurately represent functions with discontinuous first-order derivatives at the surface of the star. In the KEH and SF codes, the three-point finite-difference formulae cannot accurately represent derivatives across the discontinuous surface of the star.

The accuracy of the three codes is also estimated by the use of the two virial identities. Overall, the BGSM and SF codes show a better and more consistent agreement than the KEH code with BGSM or SF. This is largely due to the fact that the KEH code does not integrate over the whole spacetime but within a finite region around the star, which introduces some error in the computed models.

A new direct comparison of different codes is presented by Ansorg et al. [34]. Their multidomain spectral code is compared to the BGSM, KEH, and SF codes for a particular uniform density model of a rapidly rotating relativistic star. An extension of the detailed comparison in [34], which includes results obtained by the Lorene/rotstar code in [179] and by the SF code with higher resolution than the resolution used in [312]. The comparison confirms that the virial identity GRV3 is a good indicator for the accuracy of each code. The AKM code achieves nearly double-precision accuracy, while the Lorene/rotstar code has a typical relative accuracy of 10^{-4} to 10^{-6} in various quantities. The SF code at high resolution comes close to the accuracy of the Lorene/rotstar code for this model. Lower accuracies are obtained with the SF, BGSM, and KEH codes at the resolutions used in [312].

The AKM code converges to machine accuracy when a large number of about 24 expansion coefficients are used at a high computational cost. With significantly fewer expansion coefficients (and comparable computational cost to the SF code at high resolution) the achieved accuracy is comparable to the accuracy of the Lorene/rotstar and SF codes. Moreover, at the mass-shedding limit, the accuracy of the AKM code reduces to about five digits (which is still highly accurate, of course), even with 24 expansion coefficients, due to the nonanalytic behavior of the solution at the surface. Nevertheless, the AKM method represents a great achievement, as it is the first method to converge to machine accuracy when computing rapidly rotating stars in general relativity.

7.5.2 Properties of Rotating Equilibrium Stellar Structures

In the following we show a few results from the integration of Einstein's equations for a typical EoS (the nuclear EoS SLy4). As a result of the integration we get the profiles for the metric functions for given angle θ , for example. The most interesting function is the frame-dragging frequency $\omega(r, \theta)$, which decays as $1/r^3$ in the asymptotic region. The potential has a smooth behavior and attains its maximum on the rotation axis with a value of about 80 percent of the rotation of the neutron star (Fig. 7.3). It immediately decays inside the star to a value of about 20 percent on the surface of the star. This means that the absolute space has a strong differential rotation inside the star which, however, vanishes on the rotational axis.

In contrast to static stars, rotating stars form a two-parameter sequence with the central density ϵ_c as first parameter and the central rotation Ω_* as the second parameter. We only consider rigidly rotating stars. Figure 7.4 displays the total gravitational mass vs. central total energy density for evolutionary sequences of constant rest mass (constant baryon number) with increasing rotation. Above the rest mass of $2.1 M_\odot$ we find supramassive stars, which, however, can only exist as

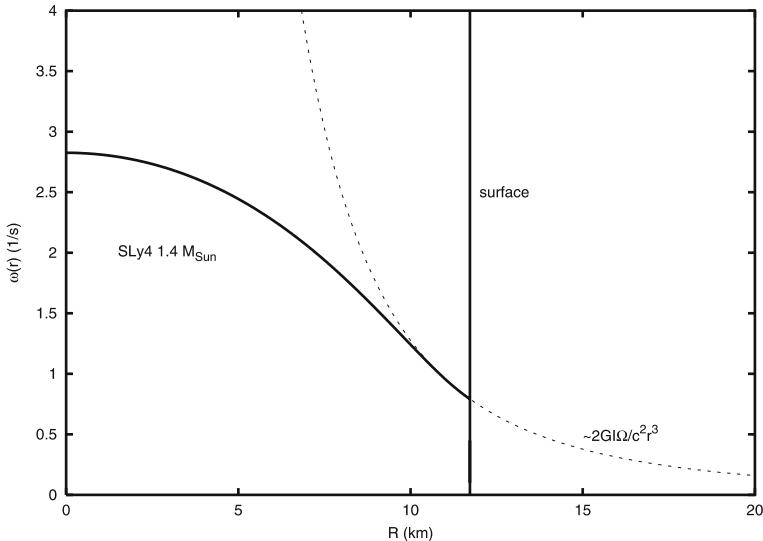


Fig. 7.3. The decay of the frame-dragging potential $\omega(r)$ as a function of the equatorial radius for the equation of state SLy4 for a star of $1.4 M_{\odot}$ rotating with a period of one second. Data provided by A. Bauswein [63]

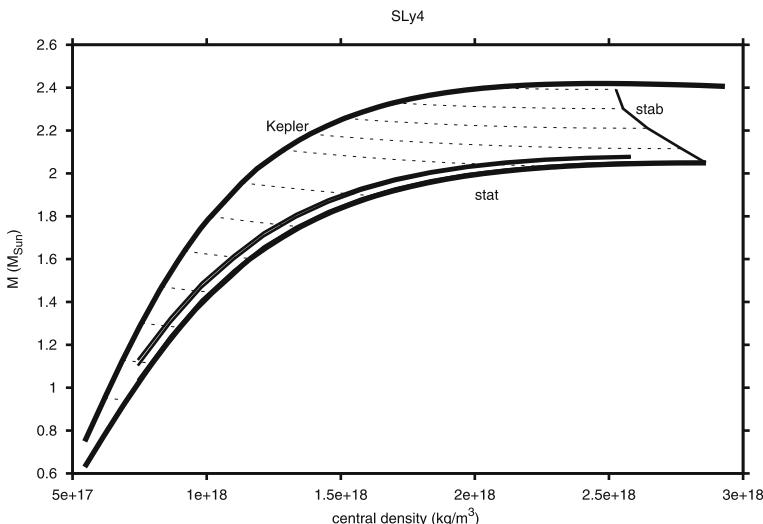


Fig. 7.4. The gravitational mass as a function of central density for SLy4 EoS. The sequences are shown for fixed rest mass M_0 ($0.6, 1.0, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.5, 2.6, 2.7, 2.8 M_{\odot}$). The *solid line* gives the nonrotating solutions, the *left line* is the mass-shed limit. For rapidly rotating stars, the upper mass limit is higher. When these stars are slowed down, they would collapse towards rotating black holes. Data have been calculated by A. Bauswein [63] using the LORENE software package with the SLy4 EoS

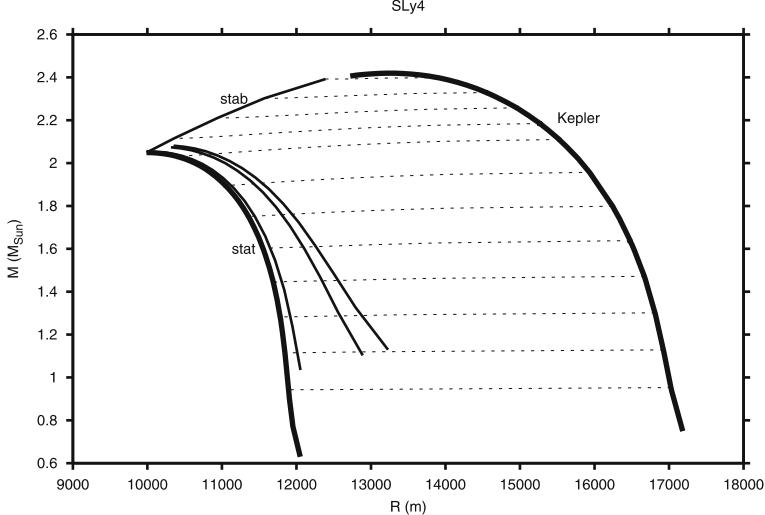


Fig. 7.5. The gravitational mass as a function of the equatorial radius. Sequences are shown for fixed rest mass M_0 with values as in the previous Figure. The *solid lines* denote sequences of constant rotation period: nonrotating, 3.4 ms, 1.56 ms, 1.4 ms and mass-shed limit (from *left* to *right*). Data have been calculated by A. Bauswein [63] using the LORENE software package with the SLy4 EoS

rotating objects. When these stars are slowed down, they would collapse to form black holes.

Figure 7.5 displays the total gravitational mass as a function of the equatorial circumferential radius for evolutionary sequences. This demonstrates that rotation inflates the equatorial bulge even for neutron star matter and leads to a strong ellipticity. This has interesting consequences for the question of the last stable orbit for millisecond pulsars. These calculations show that rotation is important for the structure of neutron stars for periods less than about 3 milliseconds.

Figure 7.6 displays the angular velocity as a function of the angular momentum. Rapidly rotating pulsars would evolve along sequences of constant baryonic mass, given by dotted lines. While normal neutron stars move towards the origin in this plot, supramassive stars would spin up, before collapsing towards a black hole. If a neutron star is sufficiently massive, the evolutionary sequence may exhibit an extended region where spin-up is allowed for.

It has been noted by many people that the maximum rotational frequency for many equations of state for neutron star matter seems to be given by a simple formula

$$\Omega_{\max} = \chi \sqrt{\frac{M_*}{M_\odot}} \left(\frac{R_e}{10 \text{ km}} \right)^{-3/2}, \quad (7.219)$$

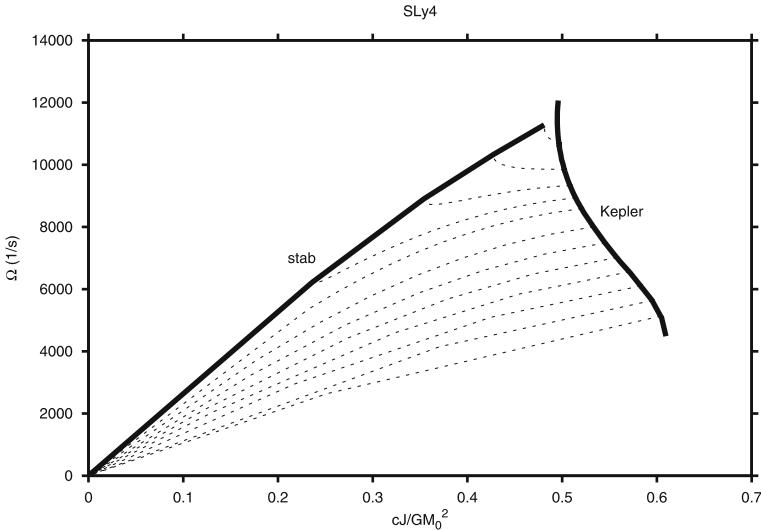


Fig. 7.6. Sequences of constant rest mass M_0 in the (Ω, J) -plane. Rapidly rotating radio pulsars would evolve along such sequences. Data have been calculated by A. Bauswein [63] using the LORENE software package with the SLy4 EoS

where $\chi \simeq 7600 s^{-1}$ and M_* and R_e are the total mass-energy and the areal radius of the maximum-mass static configuration for given EoS.

Rapidly Rotating Strange Stars

Exact models of uniformly rotating strange stars, built of self bound quark matter, are calculated within the framework of general relativity by Gourgoulhon et al. [181]. This is made possible thanks to a new numerical technique capable to handle the strong density discontinuity at the surface of these stars (BGSM). Numerical calculations are done for a simple MIT bag model equation of state of strange quark matter. Evolutionary sequences of models of rotating strange stars at constant baryon mass are calculated. Maximally rotating configurations of strange stars are determined, assuming that the rotation frequency is limited by the mass shedding and the secular instability with respect to axisymmetric perturbations. Exact formulae which give the dependence of the maximum rotation frequency, and of the maximum mass and corresponding radius of rotating configurations, on the value of the bag constant, are obtained. The values of T/W for rapidly rotating massive strange stars are significantly higher than those for ordinary neutron stars. This might indicate particular susceptibility of rapidly rotating strange stars to triaxial instabilities.

7.6 Towards Analytical Vacuum Solutions for Rotating Neutron Stars

For many purposes it would be extremely helpful to have analytic expressions for the metric coefficients at least in the exterior, vacuum region of neutron stars. Accretion theory would be much easier to be handled with analytic expressions. In fact, in the last years some new aspects have emerged in this respect using some old ideas. Essentially, one has to find out a way to combine metric functions in such a way that the resulting Einstein's equations decouple as much as possible, as in the case of the Kerr geometry (see Sect. 8.1.2).

The exterior metric of a rapidly rotating neutron star differs, however, considerably from the Kerr metric. The two metrics agree only to lowest order in the rotational velocity. At higher order, the multipole moments of the gravitational field created by a rapidly rotating compact star are different from the multipole moments of the Kerr field. There have been many attempts in the past to find analytic solutions to the Einstein equations in the stationary, axisymmetric case, that could describe a rapidly rotating neutron star. An interesting solution has been found recently by Manko et al. [267]. For nonmagnetized sources of zero net charge, the solution reduces to a three-parameter solution, involving the mass, specific angular momentum, and a parameter that depends on the quadrupole moment of the source. Although this solution depends explicitly only on the quadrupole moment, it approximates the gravitational field of a rapidly rotating star with higher nonzero multipole moments. It would be interesting to determine, whether this analytic quadrupole solution approximates the exterior field of a rapidly rotating star more accurately than the quadrupole, slow rotation approximation.

7.6.1 Weyl–Papapetrou Form

A useful procedure in this direction had already been developed by Papapetrou in 1954. The vacuum equations can be written in a more elegant way, when we choose cylindrical coordinates $\varrho = x^2$ and $z = x^3$ and use the gauge freedom of the vacuum case, $\exp(2\mu_2) = \exp(2\mu_3) \equiv \exp(2\mu)$, so that the metric assumes the form

$$ds^2 = -\exp(2\nu)c^2 dt^2 + \exp(2\psi)(d\phi - \omega dt)^2 + \exp(2\mu)(d\varrho^2 + dz^2). \quad (7.220)$$

ψ , ν , ω and μ are now functions of ϱ and z . Historically, it has been found that the vacuum equations considerably simplify, if one uses the **Weyl–Papapetrou form** of the metric

$$ds^2 = -f(c dt - \omega d\phi)^2 + \frac{1}{f} [\exp(2\gamma)(d\varrho^2 + dz^2) + \varrho^2 d\phi^2] \quad (7.221)$$

with the transformation

$$\exp(2\gamma) = \exp(2\mu) [\exp(2\nu) - \omega^2 \varrho^2 \exp(-2\nu)/c^2] \quad (7.222)$$

$$f = \exp(2\nu) - \varrho^2 \omega^2 \exp(-2\nu)/c^2 \quad (7.223)$$

$$\tilde{\omega} = \frac{\varrho^2 \omega / c}{\exp(4\nu) - \varrho^2 \omega^2 / c^2}. \quad (7.224)$$

The function f is a kind of mixture of the gravitational potential and the frame-dragging potential. $\tilde{\omega}$ is still proportional to the frame-dragging potential ω , while $\exp(2\nu)/f = \exp(2\mu)$ is now just the metric of the meridional plane. The field equations assume then the very simple form¹

$$f \nabla^2 f = \nabla f \cdot \nabla f - \frac{f^4}{\varrho^2} \nabla \tilde{\omega} \cdot \nabla \tilde{\omega} \quad (7.225)$$

$$\nabla \cdot \left(\frac{f^2}{\varrho^2} \nabla \tilde{\omega} \right) = 0, \quad (7.226)$$

and γ is determined by means of the conditions

$$4\gamma_{,\varrho} = \varrho \left[(\ln f)_{,\varrho}^2 - (\ln f)_{,z}^2 \right] \quad (7.227)$$

$$2\gamma_{,z} = \varrho (\ln f)_{,\varrho} (\ln f)_{,z}. \quad (7.228)$$

The equation for ω can now be satisfied in terms of the ansatz

$$\Psi_{,\varrho} \equiv \frac{f^2}{\varrho} \tilde{\omega}_{,z}, \quad \Psi_{,z} = -\frac{f^2}{\varrho} \tilde{\omega}_{,\varrho} \quad (7.229)$$

with the integrability condition

$$\left[\frac{\varrho}{f^2} \Psi_{,\varrho} \right]_{,\varrho} + \left[\frac{\varrho}{f^2} \Psi_{,z} \right]_{,z} = 0. \quad (7.230)$$

We can write therefore the two equations in the form of Ernst's equations [145]

$$f \nabla^2 f = \nabla f \cdot \nabla f - \nabla \Psi \cdot \nabla \Psi \quad (7.231)$$

$$f \nabla^2 \Psi = 2 \nabla f \cdot \nabla \Psi. \quad (7.232)$$

7.6.2 Ernst Equations

With the introduction of the complex *Ernst potential* (Ernst [145, 146])

$$\mathcal{E} \equiv f + i\Psi \quad (7.233)$$

we can combine the above two equations for f and Ψ into one complex equation

$$\Re[\mathcal{E}] \nabla^2 \mathcal{E} - (\nabla \mathcal{E})^2 = 0. \quad (7.234)$$

The virtue of this procedure is that the original complicated formulation of Einstein's equations in terms of four nonlinearly coupled partial differential equations has been

¹ $\nabla \cdot \mathbf{u} \equiv \frac{1}{\varrho} (\varrho u_\varrho)_{,\varrho} + u_{z,z} + \frac{1}{\varrho} u_{\phi,\phi}.$

reduced to one simple complex equation. This is a kind of a complex potential equation in the meridional plane. Sibgatullin has shown in 1991 [370] that this can be transformed to a linear integral equation with a Cauchy type kernel

$$\mathcal{E} = \int_{-1}^1 \frac{\mu(\sigma) e(\xi) d\sigma}{\sqrt{1 - \sigma^2}}, \quad (7.235)$$

where the kernel $\mu(\sigma)$ satisfies the integral equation

$$\int_{-1}^1 \frac{\mu(\sigma) [e(\xi) + e^*(\eta)] d\sigma}{(\sigma - \tau)\sqrt{1 - \sigma^2}} = 0 \quad (7.236)$$

with the definition of the complex variables

$$\xi = z + i\sigma\varrho \quad (7.237)$$

$$\eta = z + i\tau\varrho, \quad \sigma, \tau [-1, 1]. \quad (7.238)$$

The solution is determined by the values of the potential on the symmetry axis

$$e(z) \equiv \mathcal{E}(\varrho = 0, z). \quad (7.239)$$

Ernst's potential for the *Schwarzschild solution* is simply given by one parameter

$$\mathcal{E}(\varrho = 0, z) = \frac{z - M}{z + M} \quad (7.240)$$

where M denotes the total mass. This can be generalized for the Kerr solution (see later)

$$\mathcal{E}(\varrho = 0, z) = \frac{z - M + ia}{z + M + ia} \quad (7.241)$$

where now the parameter a is the spin parameter of the black hole. Mass enters as a real parameter into this potential, while spin is imaginary. One can expect that higher mass multipoles would also appear as real parameters, while spin multipoles will be imaginary. This is in fact the case.

It is by no means clear that the *result of the collapse* of a rapidly rotating stellar core will end in the Kerr solution. In fact the gravitational field for a rapidly rotating neutron star is *not* the Kerr solution, since the mass-quadrupole moment Q is an important parameter. In fact comparing with Newtonian gravity, one expects a doubly infinite series of solutions specified by the infinite number of mass-multipole moments and angular momentum multipole moments. One could try the solution for a finite multipole solution of order n in the form of (Sibgatullin and Sunyaev [371])

$$\mathcal{E}(\varrho = 0, z) = \frac{z^n - Mz^{n-1} + \sum_{j=1}^n a_j z^{n-j}}{z^n + Mz^{n-1} + \sum_{j=1}^n a_j z^{n-j}} \quad (7.242)$$

If we require that the coefficients with even index, a_{2k} , are real (corresponding to the Newtonian mass-multipole moments) and that the coefficients with odd index, a_{2k-1} , are pure imaginary (corresponding to the distribution of angular momentum), then the solution is symmetric about the equatorial plane. The entire solution on the meridional plane can then be written as

$$\mathcal{E}(\varrho, z) = \frac{\Delta_-}{\Delta_+}, \quad \Delta_\pm = \det(\mathcal{E}_{jk}^\pm) \quad (7.243)$$

for some suitable expressions \mathcal{E}_{ik} . These quantities are then found by the roots of some polynomials.

As a special case, one can consider a three-parameter solution for $n = 2$ (the so-called Manko-3 solution)

$$\mathcal{E}(\varrho = 0, z) = \frac{z^2 + (ia - M)z + Mb}{z^2 + (ia + M)z + Mb}, \quad (7.244)$$

where $a \equiv jM$ and j denotes the specific angular momentum of the source, and b is now a measure for the quadrupole mass moment of the source. This solution has been found by Manko et al. in 1994.

7.6.3 Manko's Solution

After more than 10 years of work in the field, Manko et al. [268] were finally able to find a vacuum solution involving five parameters (mass, angular momentum, charge, magnetic dipole moment and a mass quadrupole moment). This solution can be expressed in terms of rational functions. For our purpose, we neglect charge and magnetic dipole moment. If we denote by M the gravitational mass of the star, by $a = J/M$ the specific angular momentum, and introduce a parameter b which can be related to the mass quadrupole moment, then the choice for the axis values of the Ernst potential is

$$\mathcal{E}(\varrho = 0, z) = \frac{(z - M - ia)(z + ib) + d - \delta - ab}{(z + M - ia)(z + ib) + d - \delta - ab} \quad (7.245)$$

with the constants

$$\delta = \frac{-M^2 b^2}{M^2 - (a - b)^2} \quad (7.246)$$

$$d = \frac{1}{4} (M^2 - (a - b)^2). \quad (7.247)$$

Manko et al. have shown that in order to write the solution in rational form one has to introduce generalized spheroidal coordinates defined as

$$x = \frac{r_+ + r_-}{2k}, \quad y = \frac{r_+ - r_-}{2k}, \quad (7.248)$$

where $r_{\pm} = \sqrt{\varrho^2 + (z \pm k)^2}$ and $k = \sqrt{d + \delta}$. This transformation can be inverted

$$\varrho = k\sqrt{1 - y^2}\sqrt{x^2 - 1}, \quad z = kxy, \quad (7.249)$$

i.e. $|x| \geq 1$ and $-1 \leq y \leq 1$. The metric can then be written in the Weyl–Papapetrou form

$$ds^2 = -f(dt - w d\phi)^2 + \frac{k^2}{f} \left[\exp(2\gamma)(x^2 - y^2) \left(\frac{dx^2}{x^2 - 1} + \frac{dy^2}{1 - y^2} \right) + (x^2 - 1)(1 - y^2) d\phi^2 \right], \quad (7.250)$$

with the following rational ansatz

$$f = \frac{E}{D} \quad (7.251)$$

$$\exp(2\gamma) = \frac{E}{16k^8(x^2 - y^2)^4} \quad (7.252)$$

$$w = \frac{-(1 - y^2)F}{E}. \quad (7.253)$$

The functions E , F and D are polynomials in the coordinates x and y (see Stute and Camenzind [382])

$$\begin{aligned} D = & \left[4(k^2x^2 - \delta y^2)^2 + 2kmx[2k^2(x^2 - 1) + (2\delta + ab - b^2)(1 - y^2)] \right. \\ & + (a - b)[(a - b)(d - \delta) - m^2b](y^4 - 1) - 4d^2 \Big]^2 \\ & + 4y^2 \left[2k^2(x^2 - 1)[kx(a - b) - mb] - 2mb\delta(1 - y^2) \right. \\ & \left. \left. + [(a - b)(d - \delta) - m^2b](2kx + m)(1 - y^2) \right]^2 \right] \end{aligned} \quad (7.254)$$

$$\begin{aligned} E = & \left[4[k^2(x^2 - 1) + \delta(1 - y^2)]^2 \right. \\ & + (a - b)[(a - b)(d - \delta) - m^2b](1 - y^2)^2 \Big]^2 \\ & - 16k^2(x^2 - 1)(1 - y^2) \left[(a - b)[(x^2 - y^2) + 2\delta y^2] + m^2b y^2 \right]^2 \end{aligned} \quad (7.255)$$

$$\begin{aligned} F = & 8k^2(x^2 - 1) \left[(a - b)[k^2(x^2 - y^2) + 2\delta y^2] + y^2 m^2 b \right] \\ & \times \left[kmx[(2kx + m)^2 - 2y^2(2\delta + ab - b^2) - a^2 + b^2] \right. \\ & \left. - 2y^2(4\delta d - m^2 b^2) \right] \\ & + \left[4[k^2(x^2 - 1) + \delta(1 - y^2)]^2 + (a - b)[(a - b)(d - \delta) - m^2b](1 - y^2)^2 \right] \end{aligned}$$

$$\begin{aligned} & \times \{4(2kmbx + 2m^2b)[k^2(x^2 - 1) + \delta(1 - y^2)] \\ & +(1 - y^2)[(a - b)(m^2b^2 - 4\delta d) \\ & - (4kmx + 2m^2)[(a - b)(d - \delta) - m^2b]]\} . \end{aligned} \quad (7.256)$$

This metric can also be transformed to the canonical form

$$ds^2 = -\alpha^2 dt^2 + R^2(d\phi - \omega dt)^2 + \exp(2\mu_2) dx^2 + \exp(2\mu_3) dy^2 \quad (7.257)$$

with the following coefficients

$$\exp(2\mu_2) = \frac{k^2 D}{16k^8(x^2 - y^2)^4} \frac{x^2 - y^2}{x^2 - 1} \quad (7.258)$$

$$\exp(2\mu_3) = \frac{k^2 D}{16k^8(x^2 - y^2)^4} \frac{x^2 - y^2}{1 - y^2} \quad (7.259)$$

$$R^2 = \frac{k^2 D}{E} (x^2 - 1)(1 - y^2) - \frac{(1 - y^2)^2 F^2}{DE} \quad (7.260)$$

$$\omega = -\frac{fw}{R^2} \quad (7.261)$$

$$\alpha^2 = f(f - w) . \quad (7.262)$$

The quadrupole moment and the current octupole moment are given in terms of the three parameters M , a and b as (Berti and Stergioulas [70])

$$Q = -M[d - \delta - ab + a^2] \quad (7.263)$$

and

$$S_3 = -M \left[a^3 - 2a^2b + a[b^2 + 2(d - \delta)] - b(d - \delta) \right] . \quad (7.264)$$

In this approach, a and b are independent parameters. A vanishing angular momentum, $a = 0$, does not necessarily imply a vanishing Q and S_3 , as we expect on physical grounds. Indeed, a nonrotating star can have a nonvanishing quadrupole moment

$$Q(a = 0) = -\frac{M}{4} \frac{M^2 + b^2}{M^2 - b^2} \quad (7.265)$$

and a corresponding current octupole moment

$$S_3(a = 0) = -bQ(a = 0) . \quad (7.266)$$

There is no real value for b for which the quadrupole moment vanishes for a nonrotating star. This means that the solution of Manko does not reduce to the Schwarzschild solution as the rotation vanishes. Therefore, the Manko solution is not suitable for the description in the slow rotation approximation.

Berti and Stergioulas [70] have however shown that the Manko solution is quite suitable to cover rapidly rotating stars, when the rotation rate is large enough, so that the induced quadrupole deformation roughly exceeds the minimum nonvanishing oblate quadrupole deformation of the solution in the absence of rotation. Since the quadrupole moment is roughly proportional to a^2M , one expects that this analytic solution could be relevant for rotation rates satisfying $j > 0.5$, where $j = J/M^2$ is the dimensionless measure for the angular momentum of the star (this corresponds to the Kerr parameter a in the case of rotating black holes).

It is interesting that the Kerr solution is covered by Manko's solution for the case of imaginary form of the parameter b , $b = i\sqrt{M^2 - a^2}$ (Stute and Camenzind [382]). In this case, one recovers the correct expressions for the quadrupole moment, $Q = -a^2M$ and the current octupole moment $S_3 = -a^3M$.

Matching with Interior Solutions

The three parameters M , a and b are arbitrary in the analytic solution. When matched to some interior solution, only a certain combination of the parameters a and b are expected to correspond to a specific model for a neutron star. It is quite reasonable to match the analytic exterior solution by matching the gravitational mass M , the specific angular momentum a and the mass quadrupole moment Q . One can then hope to find a current octupole moment S_3 that is close to the corresponding value in the numerical model.

For this purpose, one has first to construct a highly accurate numerical model. The remaining parameter b is then determined from the numerical value of the quadrupole moment Q , given by equation (7.263) with $Q = Q_{\text{num}}$, where Q_{num} is the quadrupole moment of the numerical solution. The equation

$$-M[d(M, a, b) - \delta(M, a, b) - ab + a^2] - Q_{\text{num}} = 0 \quad (7.267)$$

has in general two possible real solutions for b , or no real solution. Thus for each set of parameters (M, a, Q_{num}) , there exist two different branches of solutions with parameters (M, a, b_-) and (M, a, b_+) , respectively (Fig. 7.7).

These two solutions correspond to two completely different spacetimes. Manko solutions are only found for a certain minimum specific angular momentum. For too high angular momentum, the κ parameter, $\kappa = \sqrt{d + \delta}$ becomes imaginary, and no Manko solution can be constructed [63]. Depending on the EoS, the Manko solution is indeed a useful approximation for neutron stars in LMXBs with rotational periods of a few milliseconds (Fig. 7.8).

Comparison with Slow Rotation

In [71], Berti et al. compare the HT slow-rotation expansion and the exact vacuum solution by Manko et al. with the full general relativistic numerical CST spacetime. They assume that the numerical models, obtained using the RNS code, are exact

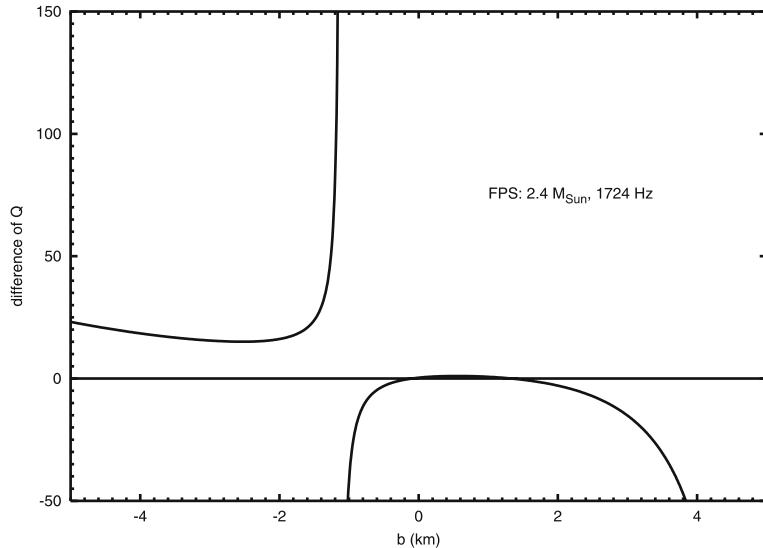


Fig. 7.7. Solutions of the equation (7.267) for a given quadrupole moment of the numerical solution. This case is for FPS and a rest mass of $2.4 M_{\odot}$ at the rotation frequency of 1724 Hz [63]

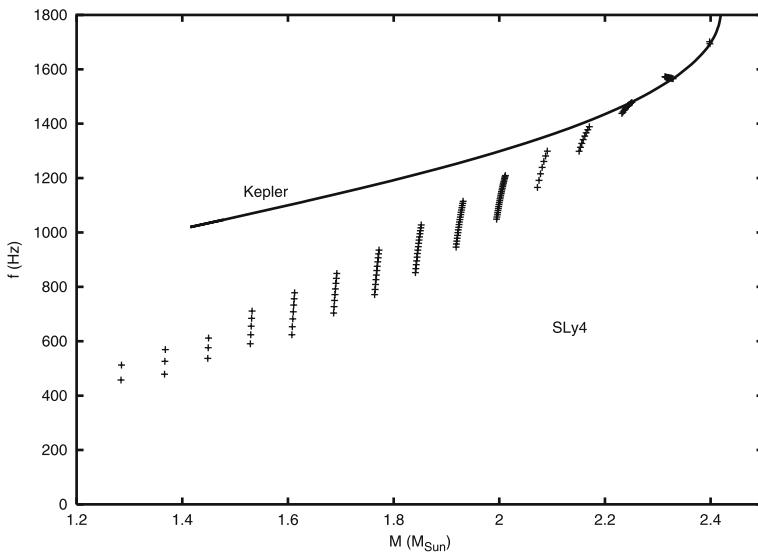


Fig. 7.8. Rotational frequency for the Manko solutions matched to internal SLy4 neutron stars [63]. Manko solutions only exist for the crosses, and not for all rapidly rotating objects. The *solid line* denotes the Kepler limit for SLy4 neutron stars

(within their numerical accuracy, that is of the order of 10^{-4} in each computed quantity). They first integrate the HT structure equations for five representative equations of state, keeping terms up to second order in the slow-rotation parameter epsilon. Then they matched these models to the CST solutions, imposing the condition that the gravitational mass and angular momentum of the models be the same. Limits of validity of the slow-rotation expansion can be estimated by computing deviations in the quadrupole moments and in the ISCO radii at different rotation rates. They found that deviations in the quadrupole moment are $\simeq 20$ percent for pulsars spinning with a period of 1.5 ms (the spin period of one of the fastest known pulsar, PSR J1939+2134). However, for these same spin rates deviations in the ISCO radii are always smaller than one percent. Since the HT approximation gives excellent predictions for ISCOs up to the fastest pulsar spin periods, it can safely be used, whenever a full numerical solution would be too cumbersome to be implemented.

7.7 On Oscillation and Formation of Rotating Neutron Stars

Finally, we mention some recent topics in the research on rapidly rotating neutron stars. Research projects center around questions such as the formation of rapidly rotating neutron stars in core collapse processes and the corresponding radiation of gravitational waves, oscillations of neutron stars and their stability in the time-dependent approach.

The study of oscillations of relativistic stars is motivated by the prospect of detecting such oscillations in electromagnetic or gravitational wave signals. In the same way that helioseismology is providing us with information about the interior of the Sun, the observational identification of oscillation frequencies of relativistic stars could constrain the high-density equation of state. The oscillations could be excited after a core collapse or during the final stages of a neutron star binary merger. Rapidly rotating relativistic stars can become unstable to the emission of gravitational waves.

In the Newtonian limit, the Lagrangian approach has been used to develop variational principles, but the Eulerian approach proved to be more suitable for numerical computations of mode frequencies and eigenfunctions [213]. A general linear perturbation of the energy density in a static and spherically symmetric relativistic star can be written as a sum of quasinormal modes that are characterized by the indices (l, m) of the spherical harmonic functions $Y_m^l(\theta, \phi)$. The Eulerian perturbation in the fluid δu^μ can be expressed in terms of vector harmonics, while the metric perturbation can be expressed in terms of spherical, vector, and tensor harmonics (see, e.g. [2]).

Oscillations of Neutron Stars

Neutron stars are laboratories for physics under extreme conditions, and may be detectable sources of gravitational waves. Many interests center around and branch out from the r -modes, fluid oscillations which feed on gravitational waves while emitting them. The signals may carry information on gravitational and nuclear physics that cannot be worked out in other ways.

Bar Mode Instability

The long-term stable evolution of rotating relativistic stars in 3D simulations has become possible through the use of High-Resolution Shock-Capturing (HRSC) methods, see Sect. 3.1. Shibata, Baumgarte, and Shapiro [366] study the dynamical bar-mode instability in differentially rotating neutron stars, in fully relativistic 3D simulations. They find that stars become unstable when rotating faster than a critical value of $T/W \simeq 0.24$. This is only somewhat smaller than the Newtonian value $T/W \simeq 0.27$.

Rapidly Rotating Core Collapse

At the end of their thermonuclear evolution massive stars have a compact central core which is mainly composed of iron and nickel. This core collapses when its mass exceeds a critical value. The collapse is brought to a halt when nuclear densities are reached. About one percent of the gravitational binding energy released during collapse will be transmitted to the stellar envelope either directly by a shock wave or indirectly by neutrinos. This causes a supernova explosion leading to the ejection of the stellar envelope and the formation of a neutron star. If no explosion occurs at all or if the explosion is too weak the stellar envelope will fall back onto the neutron star. When its mass exceeds a critical value, the neutron star will collapse to a black hole. An alternative mechanism for the collapse to a black hole is provided by phase transitions in the supranuclear equation of state of the neutron star.

Regardless of the fact that stars and especially massive stars at a late stage of their evolution are rotating, all investigations of the collapse and the explosion have assumed spherical symmetry. However, even in the case of an initially slowly rotating core, the effects of rotation can become significant. There is no efficient transport mechanism for angular momentum during the collapse. Thus, the ratio of the centrifugal force to the gravitational force grows proportional to $1/r$. During the collapse, the density increases by a factor of 10^5 , whereas the radius decreases by a factor of $\simeq 50$ and the influence of rotation increases by the same factor. In addition, the nonspherical collapse to a neutron star or a black hole is an interesting source of gravitational waves. Its detection would allow one to draw important conclusions about the dynamics of a supernova. The only other way of gaining insight into the dynamics is offered by neutrinos.

In previous investigations of the collapse of rotating cores relativistic effects have not been taken into account. However, during the collapse to a neutron star relativistic corrections of the gravitational potential can reach about 30 percent. The formation of a (rotating) black hole can only be described within a relativistic theory. In the case of rotating cores even moderate relativistic effects can be of importance. If the angular momentum of a rotating core is sufficiently large, centrifugal forces can stabilize the star below nuclear densities. This is true for the Newtonian theory as well as in the relativistic case. Due to the deeper potential in the relativistic case it is possible that the collapse is stabilized in the Newtonian

case at a density below nuclear density whereas using the correct relativistic description the same core collapses to considerably higher densities. Thus, relativistic effects can modify the collapse dynamics and thus the gravitational wave signal qualitatively.

Core collapse supernovae are a promising source of detectable gravitational waves. Most of the existing (multidimensional) numerical simulations of core collapse in general relativity have been done using approximations of the Einstein field equations. One of the most interesting such approximation is the so-called conformal flatness condition (CFC) of Isenberg, Wilson and Mathews. Building on this previous work, Cerdá-Durán et al. [110] present new results from numerical simulations of relativistic rotational core collapse in axisymmetry, aiming at improving the dynamics and the gravitational waveforms. The computer code used for these simulations evolves the coupled system of metric and fluid equations using the 3+1 formalism, specialized to a new framework for the gravitational field equations which they call CFC+. In this approach, new degrees of freedom are added to the original CFC equations, which extend them by terms of second post-Newtonian order. The corrections for CFC+ are computed solving a system of elliptic linear equations. The new formalism is assessed with time evolutions of both rotating neutron stars in equilibrium and gravitational core collapse of rotating polytropes. Gravitational wave signals for a comprehensive sample of collapse models are extracted using either the quadrupole formula or directly from the metric. They discuss the results on the dynamics and the gravitational wave emission through a detailed comparison between CFC and CFC+ simulations. The main conclusion is that, for the neutron star spacetimes analyzed in the present work, no significant differences are found among CFC, CFC+, and full general relativity, which highlights the suitability of the former.

Baiotti et al. [46] discuss the application of the Whisky code to the study of the gravitational collapse of rapidly rotating stars to Kerr black holes. The initial stellar models are modelled as relativistic polytropes which are either secularly or dynamically unstable and with uniform angular velocities which range from very slow rotation to the mass-shedding limit. They investigate the gravitational collapse by carefully studying not only the dynamics of the matter, but also that of the trapped surfaces, i.e. of both the apparent and event horizon formed during the collapse. The use of these surfaces, together with the isolated horizon framework, allows for a precise measurement of the black hole mass and spin. The ability to successfully perform these simulations for sufficiently long times relies on the possibility of excising the region of the computational domain where the singularity is formed. These authors have also found that the dynamics of the collapsing matter is strongly influenced by the initial distribution of angular momentum in the progenitor star. In particular, for initial stellar models with sufficiently high angular velocities, the collapse can lead to the formation of an unstable disk in differential rotation.

Problems

7.1. Show the equivalence of the following expressions needed for the third stellar structure equation

$$\frac{1}{\alpha} \operatorname{Div} [\alpha \nabla \psi] = \frac{1}{R} \nabla_A [R \Psi_A] + \Psi_3 \nabla_3 \mu_2 + \Psi_2 \nabla_2 \mu_3 + \frac{1}{\alpha} (\nabla_A \Psi) (\nabla_A \alpha). \quad (7.268)$$

Div is the divergence operator in three-space.

7.2. Tidal Forces and Curvature: Calculate all the components of Riemann tensor in orthonormal basis for the line element (7.8).

7.3. TOV Limit: Derive the TOV equations as the slow-rotation limit from the structure equations of rotating neutron stars.

7.4. Mechanical Equilibrium: Adapt the general relativistic hydrodynamical equations (3.32) to the metric of rotating neutron stars.

Derive the mechanical equilibrium of rotating neutron stars by starting from the 3+1 split of the hydrodynamical equations.

7.5. Angular Momentum and Quadrupole Moment: Give an estimate for the dimensionless angular momentum parameter $j = J/M^2$ for rapidly rotating neutron stars. Give also an estimate for the dimensionless quadrupole parameter $q = Q/M^3$ for neutron stars.

7.6. BGSM Equations: Show the equivalence of the structure equations derived in Sect. 7.2.3 with the equations derived by BGSM [83] for the metric

$$ds^2 = -N^2 dt^2 + B^2 r^2 \sin^2 \theta (d\phi - \omega dt)^2 + A^2 (dr^2 + r^2 d\theta^2). \quad (7.269)$$

The Einstein equations result then in a set of four elliptic equations for the metric potentials

$$\begin{aligned} \Delta_3 v &= 4\pi G A^2 (E + 3P + (E + P)U^2) \\ &\quad + \frac{B^2 r^2 \sin^2 \theta}{N^2} (\nabla \omega)^2 - (\nabla v) \cdot \nabla (v + \beta) \end{aligned} \quad (7.270)$$

$$\begin{aligned} \bar{\Delta}_3(\omega r \sin \theta) &= -16\pi G \frac{NA^2}{B} (E + P)U \\ &\quad - r \sin \theta \nabla \omega \cdot \nabla (3\beta - v) \end{aligned} \quad (7.271)$$

$$\Delta_2[(NB - 1)r \sin \theta] = 16\pi G NA^2 B Pr \sin \theta \quad (7.272)$$

$$\begin{aligned} \Delta_2(v + a) &= 8\pi G A^2 [P + (E + P)U^2] \\ &\quad + \frac{3B^2 r^2 \sin^2 \theta}{4N^2} (\nabla \omega)^2 - (\nabla v)^2. \end{aligned} \quad (7.273)$$

Here, we introduced the following abbreviations, $v = \ln N$, $a = \ln A$ and $\beta = \ln B$, and ∇ is the flat space operator. In addition, BGSM use the following differential operators

$$\Delta_3 = \partial_3^2 + \frac{2}{r} \partial_r + \frac{1}{r^2} \partial_\theta^2 + \frac{1}{r^2 \tan \theta} \partial_\theta \quad (7.274)$$

$$\tilde{\Delta}_3 = \Delta_3 - \frac{1}{r^2 \sin^2 \theta} \quad (7.275)$$

$$\Delta_2 = \partial_3^2 + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_\theta^2 \quad (7.276)$$

$$\nabla f \cdot \nabla g = (\partial_r f)(\partial_r g) + \frac{1}{r} (\partial_\theta f)(\partial_\theta g). \quad (7.277)$$

7.7. The HT Metric: Find expressions for the redshift factor α , the frame-dragging potential ω and the cylindrical radius in the HT metric.

When a static equilibrium is perturbed by rotation, the geometry of spacetime is changed. For a suitable choice of coordinates, the perturbed geometry can be expressed as follows [195]

$$\begin{aligned} ds^2 = & -\exp(\nu)[1 + 2(h_0 + h_2 P_2)] dt^2 \\ & + \frac{1 + 2(m_0 + m_2 P_2)/(r - 2M(r))}{1 - 2M(r)/r} dr^2 \\ & + r^2[1 + 2(v_2 - h_2)P_2][d\theta^2 + \sin^2 \theta (d\phi - \omega dt)^2], \end{aligned} \quad (7.278)$$

where P_2 is the Legendre polynomial of order 2. All functions are only functions of r . The surface of constant density in the rotating configuration is then given by

$$r \rightarrow r + \xi_0(r) + \xi_2(r)P_2(\cos \theta) \quad (7.279)$$

with

$$\xi_0 = -P_0^*(\epsilon + P)/(dP/dr) \quad (7.280)$$

$$\xi_2 = -P_2^*(\epsilon + P)/(dP/dr) \quad (7.281)$$

$$P_2^* = -h_2 - \frac{r^2(\omega - \Omega)^2}{3 \exp \nu}. \quad (7.282)$$

ν , $M(r)$ and $P(r)$ satisfy the unperturbed TOV equations. Derive differential equations for the perturbed quantities $h_2(r)$, $v_2(r)$, $\xi(r)$, etc. For the solutions, see Semiyoshi et al. [383].

Show that the HT metric solves the structure equations for rotating neutron stars.

7.8. Manko's Solution: By construction, Manko's solutions have two Killing fields, $k = \partial_t$ and $m = \partial_\phi$, or

$$k_\mu = (-f, f\omega, 0, 0) \quad (7.283)$$

$$m_\mu = (f\omega, \varrho^2/f - f\omega^2, 0, 0). \quad (7.284)$$

Particles move on geodesics with conserved energy E and conserved angular momentum L

$$\frac{dx^\mu}{ds} k_\mu = -E, \quad \frac{dx^\mu}{ds} m_\mu = L. \quad (7.285)$$

Discuss the solutions in the equatorial plane.

Calculate the redshift for photons emitted from the surface of a rotating star.

8 Black Holes

In this chapter we discuss the Schwarzschild solution as a global vacuum solution of Einstein's equations. We also derive the most important properties of orbital motion. This is done by using the method of effective potentials. We also show the osculating behavior of orbits near the Schwarzschild surface. Finally, the Schwarzschild solution can be extended to regions inside the Schwarzschild surface which forms an event horizon. These global vacuum solutions of Einstein's equations are called black holes. The nature of this interior vacuum of a black hole is under great debate, since a physical vacuum has some structure, which is still unknown, and is not really empty.

Whereas the solutions for rotating neutron stars can only be discussed within the framework of a numerical approach, black holes represent pure gravitational fields with a globally vanishing energy-momentum tensor, $T^{\alpha\beta} = 0$. In contrast to the exterior vacuum solutions for neutron stars, black holes are specified by the existence of an *event horizon*. This is the key feature which allows us to derive an analytic expression for the metric elements of an axisymmetric spacetime.

Einstein's general theory of relativity predicts the existence of black holes as astrophysical objects so dense that even light cannot escape from them. The boundary around the black hole, where the light cannot escape, is called the **event horizon**. In 1974, Stephen Hawking of Cambridge University theorized that a black hole is not entirely black, but could actually emit black-body, or thermal, radiation. Hawking predicted that this radiation has a well-defined temperature proportional to the gravitational force at its event horizon. In this chapter we discuss the global aspects of rotating black holes; their astrophysical relevance and the magnetospheres of black holes will be discussed in the next chapter.

8.1 The Schwarzschild Black Hole

The Schwarzschild metric is a vacuum solution of Einstein's field equations, it is valid only in the empty space outside the object. This metric can however be extended down to radii reaching the Schwarzschild surface, i.e. for $r \geq R_S$, and even towards the interior of the horizon. The mathematical form of the metric is, however, different in the object's interior. Note that the event horizon is a mathematical surface and need not coincide with any physical surface. Although the interior of a black hole,

inside the event horizon, is a region that is forever hidden from us on the outside, its properties may still be calculated.

The Schwarzschild solution

$$ds^2 = - \left(1 - \frac{2GM}{c^2r}\right) c^2 dt^2 + \left(1 - \frac{2GM}{c^2r}\right)^{-1} dr^2 + r^2 d\Omega^2 \quad (8.1)$$

can be considered as a global solution of the vacuum equations, $T_{\mu\nu} = 0$, i.e. $R_{\mu\nu} = 0$ everywhere, even at $r = 0$. In this case, the Schwarzschild solution is called a **black hole**, since nothing can escape from the horizon at $r = R_S$. One can think of the horizon as the place where the escape velocity equals the velocity of light. Outside of the horizon, the escape velocity is less than the speed of light, so if you fire your rockets hard enough, you can give yourself enough energy to get away. But if you find yourself inside the horizon, then no matter how powerful your rockets are, you cannot escape. Incidentally, the name “black hole” was invented by John Archibald Wheeler, and seems to have stuck because it was much catchier than previous names. Before Wheeler came along, these objects were often referred to as “frozen stars.”

A black hole is a region of an asymptotically flat spacetime from which nothing can escape, not even light.

The light cones distort as $r \rightarrow 2GM/c^2$, so that no future-directed time-like or null worldline can reach $r > 2GM/c^2$ from $r \leq 2GM/c^2$. No signal from a star’s surface can escape to infinity, once the surface has passed through $r = 2GM/c^2$. The star has collapsed to a black hole. For an external observer, the surface never actually reaches $r = 2GM/c^2$, but as $r \rightarrow 2GM/c^2$ the redshift of light leaving the surface increases exponentially fast and the star disappears from view within a time GM/c^3 . The late time appearance is then dominated by photons escaping from the unstable photon orbit $r = 3GM/c^2$. If an observer stayed outside a black hole, while an intrepid observational general relativist dove into the black hole, sending back signals all the time, he would simply see the signals reach him more and more slowly. This is a consequence of our discussion of the gravitational redshift. As infalling astronauts approach $r = 2GM/c^2$, any fixed interval $\Delta\tau_1$ of their proper time corresponds to a longer and longer interval from the outside point of view. This continues forever; we would never see astronauts cross $r = 2GM/c^2$, we would just see them move more and more slowly.

8.1.1 Tortoise Coordinates and Null Cones

The problem with our current coordinates is that $dt/dr \rightarrow \infty$ along radial null geodesics which approach $r = 2GM/c^2$; progress in the r -direction becomes slower and slower with respect to the coordinate time t . We can try to fix this problem by replacing t with a coordinate which “moves more slowly” along null geodesics. We may explicitly solve the condition characterizing radial null curves

$$dt = \pm \frac{1}{\sqrt{1 - 2GM/r}} dr = \pm dr_* \quad (8.2)$$

to obtain

$$t = \pm r_* + \text{const}, \quad (8.3)$$

where the **tortoise or Regge–Wheeler coordinate** r_* is defined by

$$r_* = r + \frac{2GM}{c^2} \ln \left(\frac{c^2 r}{2GM} - 1 \right). \quad (8.4)$$

As r ranges from $2M$ to infinity, r_* ranges from $-\infty$ to $+\infty$. In terms of the tortoise coordinate, the Schwarzschild metric is written in ingoing Eddington–Finkelstein coordinates¹

$$\begin{aligned} ds^2 &= \left(1 - \frac{2GM}{r}\right) (-dt^2 + dr_*^2) + r^2 d\Omega^2 \\ &= -\left(1 - \frac{2GM}{r}\right) dv^2 + 2drdv + r^2 d\Omega^2. \end{aligned} \quad (8.5)$$

This represents some progress, since the light cones now behave regularly at $r = 2M$ (Fig. 8.1); furthermore, none of the metric coefficients becomes infinite at $r = 2GM$ (although both g_{tt} and $g_{r_*r_*}$ become zero). The price we pay, however, is that the surface of interest at $r = 2M$ has just been pushed to infinity.

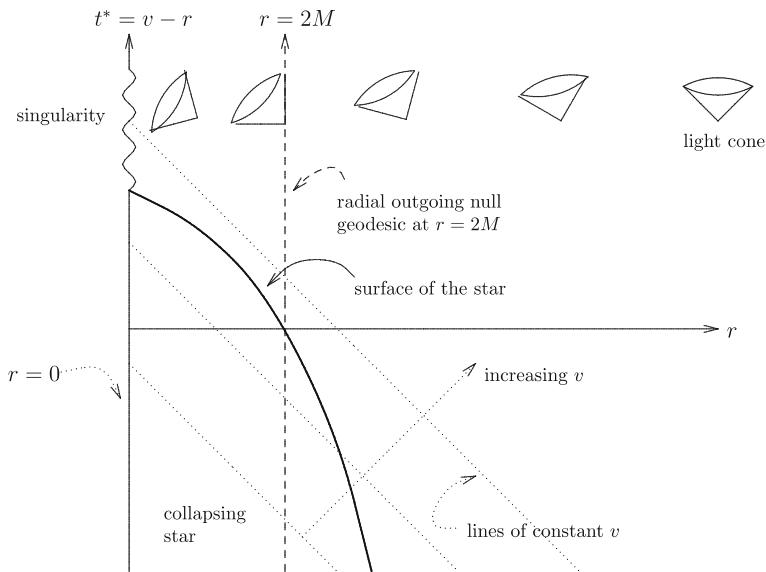


Fig. 8.1. Finkelstein diagram of a collapsing star, which is a plot of $t^* = v - r$ against r . The null cones, which are Minkowski-like at infinity, are tilted near the central object, so that no future-directed time-like or null worldline can reach $r > 2GM/c^2$ from $r \leq 2GM/c^2$

¹ Just compute dr_* and plug in to show that Schwarzschild follows.

This metric is initially defined only for $r > 2M$, since the relation $v = t + r_*(r)$ between v and r is only defined for $r > 2M$. It can however be extended analytically to all values $r > 0$ by means of the relation

$$r_* = r + 2M \ln \left| \frac{r}{2M} - 1 \right|. \quad (8.6)$$

Since the cross-term $dr dv$ is nonsingular at $r = 2M$, this singularity in the Schwarzschild metric was really only a coordinate singularity.

8.1.2 Roads towards Black Hole Formation

The formation of stellar black holes is linked up with the evolution of massive stars. At the peak of their lifetime, a star loses its fuel. It stops from burning hydrogen into helium, because it consumed all its helium. During the end of its life it started to fuse helium into other heavier elements (the heaviest one being Fe) which reached the inner part of the star. That is how basically all heavy elements in our Universe formed. But then, at the very end of its life, as waves of energy eject from the star's core which is starting to collapse, nucleosynthesis takes place, that is the star fuses its atmosphere of helium with the iron core. So, all these heavy elements and millions of neutrinos are ejected in waves as the star breaks apart. This explosion of a star is called supernova. Of course, it depends on the mass of the star, whether there is a remnant left out of the huge explosion or not. Now the star can have two roads to follow depending on the mass of their remnants. If the remnant of the star is less than $1.44 M_\odot$ (the Chandrasekhar limit), the star will become a white dwarf, that is a dead star which still has some kind of internal energy so it will still have some luminosity, but the star is fading with time.

If the remnant is roughly $1.4 M_\odot$, the core will collapse. The atom's protons and electrons will merge together and their particles will recombine to form neutrons. The result will be a new star made entirely of neutrons with a thin crust made of heavy nuclei. This is a neutron star. The structure of these stars was discussed in Chap. 6.

The other road a star can take is if its core is greater than about two solar masses. Then, we can finally talk about black holes. The pressure of the collapse and the gravity force of the entire mass squashing the core condenses all the matter together. In this case, the neutrons are not only tightly packed together, but literally forced together. If before the neutrons could stand being put together, now the pressure is too high for them resisting these forces. The result is a further collapse. If the star can collapse itself smaller than a certain radius, then it becomes a black hole with a radius given by the Schwarzschild radius.

Supermassive black holes (with masses of millions to billions of times the mass of the Sun) are found in the centers of most, if not all, massive galaxies, and the black hole masses scale with the galaxy masses, so that larger black holes reside in larger galaxies. Black holes probably evolve as material, such as gas, dust, stars and even other black holes, gets sucked in by the strong gravitational pull. The black hole seeds originally may have formed from the explosions of the first stars

or from the collapse of clumps of gas in the early Universe. Each of these different formation scenarios leads to very different numbers of intermediate-mass black holes left over in the Universe today. Gravitational wave experiments, especially the Laser Interferometric Space Antenna (LISA) expect to be very sensitive to the merging of 100,000-solar-mass black holes.

8.1.3 The Kruskal Extension

The fact that we never see the infalling astronauts reach $r = 2GM/c^2$ is a meaningful statement, but the fact that their trajectory in the $t - r$ plane never reaches there is not. It is highly dependent on our coordinate system, and we would like to ask a more coordinate-independent question (such as, do the astronauts reach this radius in a finite amount of their proper time?). The best way to do this is to change coordinates to a system which is better behaved at $r = 2GM/c^2$. There does exist a set of such coordinates.

Eddington–Finkelstein Coordinates

We can now try new coordinates u and v by means of (these are the *Eddington–Finkelstein coordinates*)²

$$u = t - r_* , \quad v = t + r_* , \quad (8.7)$$

or

$$v - u = 2r_* , \quad u + v = 2t . \quad (8.8)$$

Then

$$dt^2 = \frac{1}{4} [dv^2 + du^2 + 2du\,dv] \quad (8.9)$$

and

$$dr^2 = \left(1 - \frac{2GM}{r}\right)^2 dr_*^2 = \frac{1}{4} \left(1 - \frac{2GM}{r}\right)^2 [dv^2 + du^2 - 2du\,dv] . \quad (8.10)$$

This implies

$$ds^2 = - \left(1 - \frac{2GM}{r}\right) du\,dv + r^2 d\Omega^2 . \quad (8.11)$$

We can now extend the definition of the (u, v) -coordinates to the interior region by replacing the expression for r_* in equation (8.4) by means of

$$r_* = r + 2M \ln \left| \frac{r}{2M} - 1 \right| . \quad (8.12)$$

² In the following, we often use units with $G = 1 = c$.

We use absolute bars around the argument of the logarithm, so that r_* is now defined inside and outside the horizon at $r = 2M$. By plotting lines of u and $v = \text{constant}$ both inside and outside the horizon, one can see that each point with $r \neq 0$ and $r \neq 2M$ can be uniquely identified by the intersection of null geodesics. Hence, (u, v) is a viable coordinate system except for $r = 0$ and $r = 2M$. The coordinate r is implicitly defined as a function of u and v

$$\frac{v - u}{4M} = \frac{r_*(r)}{2M} = \frac{r}{2M} + \ln \left| \frac{r}{2M} - 1 \right|. \quad (8.13)$$

This implies the metric expression

$$ds^2 = \mp \frac{2M \exp(-r/2M)}{r} \exp((v - u)/4M) du dv. \quad (8.14)$$

Kruskal Coordinates and Maximal Extension

This representation of the metric is not an improvement over the Schwarzschild formulation, since the system is discontinuous across the horizon. However, the metric coefficients themselves approach finite values there. The discontinuity is therefore finite and in some sense less serious than a coordinate singularity. In fact, it is trivial to get rid of the discontinuity by means of the transformation to new coordinates U and V defined as follows (Kruskal–Szekeres (KS) coordinates)

$$U = \exp(u/4M) = \exp((r + t)/4M) \sqrt{\frac{r}{2M} - 1} \quad (8.15)$$

$$V = -\exp(-v/4M) = -\exp[(r - t)/4M] \sqrt{\frac{r}{2M} - 1}. \quad (8.16)$$

Then

$$\begin{aligned} dU dV &= \frac{1}{(4M)^2} \exp((u - v)/4M) du dv \\ &= \frac{1}{(4M)^2} \exp(r/2M) \left(\frac{r}{2M} - 1 \right) du dv, \end{aligned} \quad (8.17)$$

in terms of which the metric becomes

$$ds^2 = -\frac{32M^3}{r} \exp(-r/2M) dU dV + r^2 d\Omega^2,$$

(8.18)

where $r(U, V)$ is given implicitly by $UV = -\exp(r_*/2M)$ or

$$UV = -\left(\frac{r}{2M} - 1 \right) \exp(r/2M). \quad (8.19)$$

In terms of these coordinates, the metric is now totally regular at $r = 2M$. The curves of constant U and V are null geodesics. The singularity at $r = 0$ corresponds to $UV = 1$.

The Kruskal coordinates (T, R) are then linear combinations defined as

$$T = \frac{1}{2}(U + V) = \sqrt{\frac{r}{2M} - 1} \exp(r/4M) \sinh(t/4M) \quad (8.20)$$

$$R = \frac{1}{2}(V - U) = \sqrt{\frac{r}{2M} - 1} \exp(r/4M) \cosh(t/4M) \quad (8.21)$$

with the metric given as

$$ds^2 = -\frac{32M^3}{r} \exp(-r/2M) (dT^2 - dR^2) + r^2 d\Omega^2. \quad (8.22)$$

r is defined implicitly from

$$R^2 - T^2 = \left(\frac{r}{2M} - 1\right) \exp(r/2M). \quad (8.23)$$

The coordinates (T, R, θ, ϕ) are known as **Kruskal coordinates**, or sometimes Kruskal–Szekeres coordinates. Note that T is the time-like coordinate. The two-space (T, R) is conformally flat, i.e. the light cones are as in Minkowski space (Fig. 8.2). The coordinate transformation from (t, r) to (T, R) is implicit, so we cannot find $r = r(T, R)$, for example. One has, however, the following relations

$$UV = T^2 - R^2 = -\exp(r/2M) \left(\frac{r}{2M} - 1\right) \quad (8.24)$$

$$\frac{U}{V} = \frac{T - R}{T + R} = -\text{sgn}(f) \exp(-\text{sgn}(V)t/2M), \quad (8.25)$$

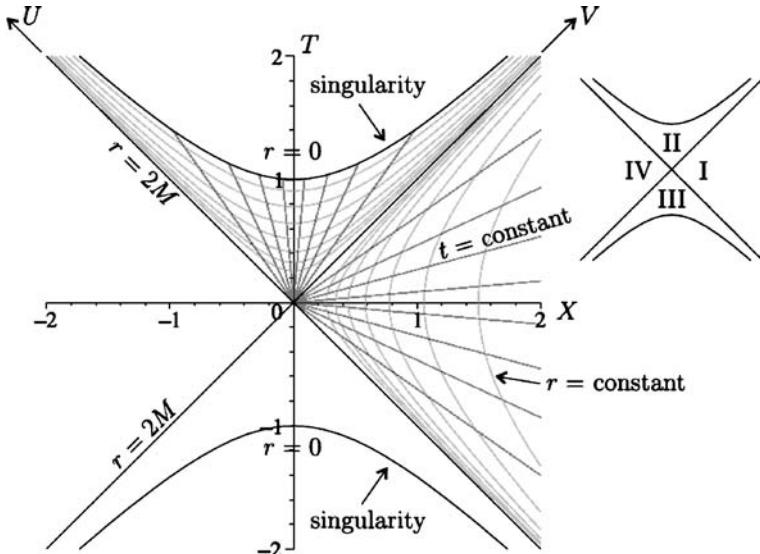


Fig. 8.2. Kruskal diagram: each point on the diagram is a two-sphere. Surfaces of constant t (straight lines) and r (hyperbola) are drawn according to the original coordinate transformation. Schwarzschild coordinates only cover regions I and II of the extended manifold

where $f = 1 - 2M/r$ is the Schwarzschild factor. The singularity at $r = 0$ is described by $UV = 1$ hyperbola.

Our original coordinates (t, r) were only well defined for $r > 2M$, which is only a part of the manifold portrayed on the Kruskal diagram (Fig. 8.2). It is convenient to divide the diagram into four regions.

The original region is region I; by following future-directed null rays we reach region II, and by following past-directed null rays we reach region III. If we had explored space-like geodesics, we would have been led to region IV. The definitions which relate (R, T) to (t, r) are really only mathematically correct in region I; in the other regions it is necessary to introduce appropriate minus signs to prevent the coordinates from becoming imaginary.

The Kruskal coordinates have a number of miraculous properties. Like the $(t, r*)$ coordinates, the radial null curves look like they do in flat space

$$T = \pm R + \text{const.} \quad (8.26)$$

Unlike the $(t, r*)$ coordinates, however, the event horizon $r = 2GM$ is not infinitely far away; in fact it is defined by

$$T = \pm R. \quad (8.27)$$

More generally, we can consider the surfaces $r = \text{const}$. These satisfy

$$R^2 - T^2 = \text{const.} \quad (8.28)$$

Thus, they appear as hyperbolae in the T - R plane. Furthermore, the surfaces of constant t are given by

$$\frac{R}{T} = \tanh(t/4GM). \quad (8.29)$$

The coordinates (T, R) should be allowed to range over every value they can take without hitting the real singularity at $r = 0$; the allowed region is therefore $-\infty < R < \infty$ and $T^2 < R^2 + 1$. We can now draw a spacetime diagram in the T - R plane (with θ and ϕ suppressed), known as a **Kruskal diagram** (Fig. 8.2), which represents the entire spacetime corresponding to the Schwarzschild metric.

The Kruskal extension is a remarkable new spacetime. Region II is what we think of as the black hole interior. Once anything travels from region I into II, it can never return. In fact, every future-directed path in region II ends up hitting the singularity at $r = 0$. Not only can you not escape back to region I, you cannot even stop yourself from moving in the direction of decreasing r , since this is simply the time-like direction. As a spaceship falls towards the singularity, its front and back will be pulled apart from each other, while its body is squeezed to infinitesimal thinness.

Regions III and IV might be somewhat unexpected. Region III is simply the time-reverse of region II, a part of spacetime from which things can escape to us, while we can never get there. It can be thought of as a “white hole.” There is a singularity

in the past, out of which the Universe appears to spring. The boundary of region III is sometimes called the past event horizon, while the boundary of region II is called the future event horizon. Region IV, meanwhile, cannot be reached from our region I either forward or backward in time (nor can anybody from over there reach us). It is another asymptotically flat region of spacetime, a mirror image of ours. It can be thought of as being connected to region I by a “wormhole,” a neck-like configuration joining two distinct regions.

8.1.4 Penrose Diagram – the Conformal Structure of Infinity

We have seen that the Kruskal coordinate system provides a very useful representation of the Schwarzschild geometry. Before moving on to other types of black holes, we will introduce one more way of thinking about this spacetime, the *Penrose (or Carter–Penrose, or conformal) diagram*. The idea is to do a conformal transformation which brings the entire manifold onto a compact region such that we can fit the spacetime on a piece of paper.

Minkowski Space

Let us begin with Minkowski space in polar coordinates, to see how the technique works. Technically the worldline $r = 0$ represents a coordinate singularity and should be covered by a different patch, but we all know what is going on so we will just assume that $r = 0$ is well-behaved. Then we switch to null coordinates

$$u = (t + r)/2, \quad v = (t - r)/2, \quad (8.30)$$

with the ranges $\infty < u, v < \infty$ and $v \leq u$. The metric in these coordinates is given by

$$ds^2 = -2(du\,dv + dv\,du) + (u - v)^2 d\Omega^2. \quad (8.31)$$

We now want to change to coordinates in which “infinity” takes on a finite coordinate value. A good choice is

$$U \equiv \arctan(u) \quad V \equiv \arctan(v), \quad (8.32)$$

with the ranges $-\pi/2 \leq U, V \leq \pi/2$ and $V \leq U$. By using

$$dU = \frac{du}{1 + u^2} \quad (8.33)$$

and

$$\cos(\arctan(u)) = \frac{1}{\sqrt{1 + u^2}} \quad (8.34)$$

we are lead to

$$du \, dv + dv \, du = \frac{1}{\cos^2(U) \cos^2(V)} (dU \, dV + dV \, dU) \quad (8.35)$$

and

$$\begin{aligned} (u - v)^2 &= (\tan U - \tan V)^2 \\ &= \frac{1}{\cos^2(U) \cos^2(V)} (\sin U \cos V - \cos V \sin U)^2 \\ &= \frac{1}{\cos^2(U) \cos^2(V)} \sin^2(U - V). \end{aligned} \quad (8.36)$$

Therefore, the Minkowski metric in these coordinates is

$$ds_M^2 = -\frac{1}{\cos^2(U) \cos^2(V)} [2(dU \, dV + dV \, dU) - \sin^2(U - V)^2 d\Omega^2]. \quad (8.37)$$

This has a certain appeal, since the metric appears as a fairly simple expression multiplied by an overall factor. We can make it even better by transforming back to a time-like coordinate η and a space-like (radial) coordinate χ , via

$$\eta \equiv U + V, \quad \chi \equiv U - V \quad (8.38)$$

with ranges $-\pi < \eta < +\pi$ and $0 \leq \chi < \pi$. In these coordinates, the metric is

$$ds_M^2 = w^{-2} (-d\eta^2 + d\chi^2 + \sin^2 \chi d\Omega^2), \quad (8.39)$$

where $w = \cos U \cos V = (\cos \eta + \cos \chi)/2$. The Minkowski metric may therefore be thought of as related by a conformal transformation to the unphysical metric

$$ds^2 = w^2 ds_M^2 = -d\eta^2 + d\chi^2 + \sin^2 \chi d\Omega^2. \quad (8.40)$$

This describes the manifold $R \times S^3$, where the three-sphere is maximally symmetric and static. There is curvature in this metric, and it is not a solution to the vacuum Einstein's equations. This should not bother us, since it is unphysical; the true physical metric, obtained by a conformal transformation, is simply flat space-time. In fact this metric is that of the **Einstein static Universe**, a static (but unstable) solution to Einstein's equations with a perfect fluid and a cosmological constant. Of course, the full range of coordinates on $R \times S^3$ would usually be $-\infty < \eta < +\infty$, $0 \leq \chi \leq \pi$, while Minkowski space is mapped into the subspace. The entire $R \times S^3$ can be drawn as a cylinder, in which each circle is a three-sphere, as shown in Fig. 8.3.

The shaded region represents Minkowski space. Note that each point (η, χ) on this cylinder is half of a two-sphere, where the other half is the point $(\eta, -\chi)$. We can unroll the shaded region to portray Minkowski space as a triangle, as shown in Fig. 8.4.

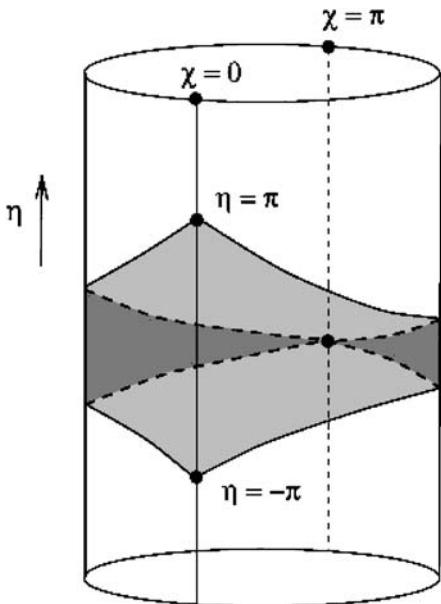


Fig. 8.3. The Einstein static Universe represented by a cylinder. Each *point* represents one half of a two-sphere. The *shaded* region is conformal to Minkowski spacetime. Its boundary may be regarded as conformal infinity of Minkowski spacetime

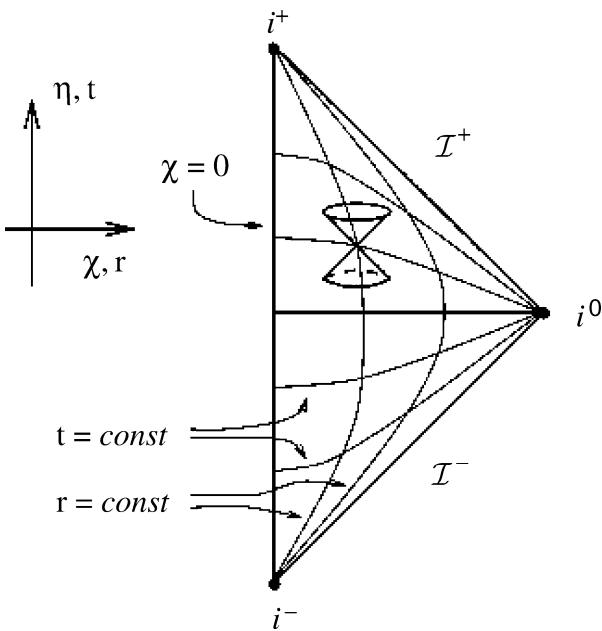


Fig. 8.4. Penrose diagram for Minkowski space

This is the **Penrose diagram for the Minkowski space**. Each point represents a two-sphere. In fact Minkowski space is only the interior of the above diagram (including $\chi = 0$); the boundaries are not part of the original spacetime. Together they are referred to as **conformal infinity**. The structure of the Penrose diagram allows us to subdivide conformal infinity into a few different regions:

- i^+ : future time-like infinity ($\eta = \pi, \chi = 0$);
- i^0 : spatial infinity ($\eta = 0, \chi = \pi$);
- i^- : past time-like infinity ($\eta = -\pi, \chi = 0$);
- \mathcal{I}^+ : future null infinity ($\eta = \pi - \chi, 0 < \chi < \pi$);
- \mathcal{I}^- : past null infinity ($\eta = -\pi + \chi, 0 < \chi < \pi$).

There are a number of important features of the Penrose diagram for Minkowski spacetime. The points i^+ , and i^- can be thought of as the limits of space-like surfaces whose normals are time-like; conversely, i^0 can be thought of as the limit of time-like surfaces whose normals are space-like. Radial null geodesics are at ± 45 degrees in the diagram. All time-like geodesics begin at i^- and end at i^+ ; all null geodesics begin at \mathcal{I}^- and end at \mathcal{I}^+ ; all space-like geodesics both begin and end at i^0 . On the other hand, there can be nongeodesic time-like curves that end at null infinity (if they become asymptotically null).

It is nice to be able to fit all of Minkowski space on a small piece of paper, but we do not really learn much that we did not already know. Penrose diagrams are more useful when we want to represent slightly more interesting spacetimes, such as those for black holes. The original use of Penrose diagrams was to compare spacetimes to Minkowski space “at infinity” – the rigorous definition of *asymptotically flat* is basically that a spacetime has a conformal infinity just like Minkowski space. We will not pursue these issues in detail, but instead turn directly to analysis of the Penrose diagram for a Schwarzschild black hole.

Penrose Diagram for Kruskal

We will not go through the necessary manipulations in detail, since they parallel the Minkowski case with considerable additional algebraic complexity. The (u'', v'') part of the metric (that is, at constant angular coordinates) is now conformally related to Minkowski space. In the new coordinates the singularities at $r = 0$ are straight lines that stretch from time-like infinity in one asymptotic region to time-like infinity in the other. The Penrose diagram for the maximally extended Schwarzschild solution thus looks like Fig. 8.5.

The only real subtlety about this diagram is the necessity to understand that i^+ and i^- are distinct from $r = 0$ (there are plenty of time-like paths that do not hit the singularity). Notice also that the structure of conformal infinity is just like that of Minkowski space, consistent with the claim that Schwarzschild is asymptotically flat. Also, the Penrose diagram for a collapsing star that forms a black hole is what you might expect.

Once again the Penrose diagrams for these spacetimes don’t really tell us anything we didn’t already know; their usefulness will become evident when we consider more

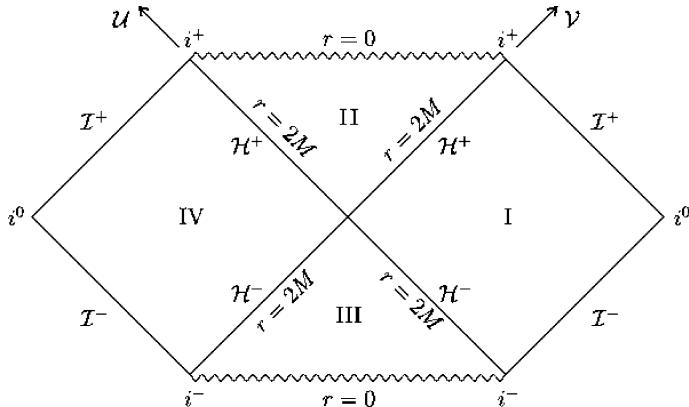


Fig. 8.5. Penrose diagram for the maximally extended Schwarzschild spacetime. Two lines drawn at 45 deg angles should intersect in the diagram only if the corresponding two light rays intersect in the actual spacetime. A Penrose diagram can be used as a concise illustration of spacetime regions that are accessible to observation. The horizontal boundary lines of a Penrose diagram correspond to the “infinity” or to singularities where light rays must end. Thus, Penrose diagrams are also useful in the study of asymptotic properties of spacetimes and singularities

general black holes. In principle there could be a wide variety of types of black holes, depending on the process by which they were formed. Surprisingly, however, this turns out not to be the case; no matter how a black hole is formed, it settles down (fairly quickly) into a state which is characterized only by the mass, charge, and angular momentum. This property, which must be demonstrated individually for the various types of fields which one might imagine going into the construction of the hole, is often stated as **black holes have no hair**. One can demonstrate, for example, that a hole which is formed from an initially inhomogeneous collapse shakes off any lumpiness by emitting gravitational radiation. This is an example of a “no-hair theorem.” If we are interested in the form of the black hole after it has settled down, we thus need only to concern ourselves with charged and rotating holes. In both cases there exist exact solutions for the metric, which we can examine closely.

A Penrose diagram (named for mathematical physicist Roger Penrose) is a two-dimensional diagram that captures the causal relations between different points in spacetime. It is an extension of a Minkowski diagram where the vertical dimension represents time, and the horizontal dimension represents space, and slanted lines at an angle of 45 degrees correspond to light rays. The biggest difference is that locally, the metric on a Penrose diagram is conformally equivalent to the actual metric in spacetime. The conformal factor is chosen such that the entire infinite spacetime is transformed into a Penrose diagram of finite size. For spherically symmetric spacetimes, every point in the diagram corresponds to a two-sphere (Fig. 8.5).

We have seen how the maximally extended Schwarzschild geometry involves several distinct regions, each of which can be covered by a (t, r) coordinate patch. We now are interested in the causal links between these different patches; can an

event in region II, for example, influence those in regions III? For this it is useful to introduce the following rescaling

$$\mathcal{U} = \frac{2}{\pi} \arctan U, \quad \mathcal{V} = \frac{2}{\pi} \arctan V. \quad (8.41)$$

This transformation maps the real plane spanned by (U, V) into the finite area $\mathbf{I} = [-1, 1] \times [-1, 1]$, i.e. instead of an infinite two-dimensional plane we have only to consider a finite square (Fig. 8.5). In these coordinates, the line element is given by

$$ds^2 = -\frac{8\pi^2 M^3 \exp(-r/2M)}{r} \sec^2\left(\frac{\pi \mathcal{U}}{2}\right) \sec^2\left(\frac{\pi \mathcal{V}}{2}\right) d\mathcal{U} d\mathcal{V}. \quad (8.42)$$

This line element has coordinate singularities as \mathcal{U} or \mathcal{V} approach ± 1 , but this is of no major importance. This metric is in fact identical to the Kruskal–Szekeres form (8.18), except for a multiplicative prefactor. Metrics of this type are called conformally identical. They have the properties that their null geodesics are the same. While Penrose diagrams share the same basic coordinate vector system of other spacetime diagrams for local asymptotically flat spacetime, it introduces a system of representing distant spacetime by shrinking or “crunching” distances that are further away. Straight lines of constant time and space coordinates therefore become hyperbolae, which appear to converge at points in the corners of the diagram. These points represent “conformal infinity” for space and time (represented as i -points in Fig. 8.5).

The key features of a Penrose diagram are given in Table 8.1.

We can now investigate which special locations in the Kruskal diagram get mapped in the new coordinate system. In the Kruskal system, the **future horizon** was defined as $U = 0$ for $V > 0$ and $V = 0$ for $U > 0$, which now translates into (see Fig. 8.5)

$$\mathcal{H}^+ = \{(\mathcal{U}, \mathcal{V}) \in \mathbf{I} : \mathcal{U} = 0 \text{ for } \mathcal{V} > 0, \mathcal{V} = 0 \text{ for } \mathcal{U} > 0\}. \quad (8.43)$$

Similarly, the **past horizon** is given by

$$\mathcal{H}^- = \{(\mathcal{U}, \mathcal{V}) \in \mathbf{I} : \mathcal{U} = 0 \text{ for } \mathcal{V} < 0, \mathcal{V} = 0 \text{ for } \mathcal{U} < 0\}. \quad (8.44)$$

Table 8.1. Special features in the Penrose diagram

Feature	r	t	u	v	\mathcal{U}	\mathcal{V}
Future singularity	0				$1 - \mathcal{V}$	$1 - \mathcal{U}$
Past singularity	0				$-1 - \mathcal{V}$	$-1 - \mathcal{U}$
Future horizon \mathcal{H}^+	$2M$	$+\infty$	$+\infty$	Finite	$0, \mathcal{V} > 0$	$0, \mathcal{U} > 0$
Past horizon \mathcal{H}^-	$2M$	$-\infty$	finite	$-\infty$	$0, \mathcal{V} < 0$	$0, \mathcal{U} < 0$
Future time-like infinity i^+	$> 2M$	$+\infty$	$+\infty$	$+\infty$	$0, +1$	$+1, \text{or } 0$
Past time-like infinity i^-	$> 2M$	$-\infty$	$-\infty$	$-\infty$	$0, \text{or } -1$	$-1, \text{or } 0$
Space-like infinity i^0	$+\infty$	finite	$-\infty$	$+\infty$	± 1	∓ 1
Future null infinity \mathcal{I}^+	$+\infty$	$+\infty$	finite	$+\infty$	$\mathcal{U} = +1$	$\mathcal{V} = +1$
Past null infinity \mathcal{I}^-	$-\infty$	$-\infty$	$-\infty$	Finite	$\mathcal{U} = -1$	$\mathcal{V} = -1$

Wormholes

One of the most intriguing aspect of the maximally extended Schwarzschild geometry is the existence of two asymptotically flat regions I and IV. In a way, each of these represent distinct worlds, with separate observers and politics. The two parallel Universes are clearly causally disconnected from each other, it is impossible to travel from one to the other without travelling faster than light. Although this may seem more the stuff of science fiction than science fact, physicists first dreamed up the idea of wormholes. In 1935, Albert Einstein and Nathan Rosen realized that general relativity allows the existence of “bridges,” originally called Einstein–Rosen bridges, but now known as wormholes. These spacetime tubes act as shortcuts connecting distant regions of spacetime. By journeying through a wormhole, you could travel between the two regions faster than a beam of light would be able to if it moved through normal spacetime. As with any mode of faster-than-light travel, wormholes offer the possibility of time travel.

8.2 Geodetic Motions in Schwarzschild Spacetime

As we have seen earlier, the equations governing the geodesics in a spacetime can be derived from the energy integral given by the Lagrangian

$$2\mathcal{L} = -g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}, \quad (8.45)$$

where λ is some affine parameter along the geodesic. For time-like geodesics, λ may be identified with the proper time τ .

8.2.1 A Lagrangian

For Schwarzschild the Lagrangian is (here we work in units $c = 1 = G$)

$$\mathcal{L} = \frac{1}{2} \left[\left(1 - \frac{2M}{r}\right) i^2 - \frac{\dot{r}^2}{1 - 2M/r} - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right], \quad (8.46)$$

where a dot denotes differentiation with respect to λ . The corresponding canonical momenta are

$$p_i = \frac{\partial \mathcal{L}}{\partial i} = \left(1 - \frac{2M}{r}\right) i \quad (8.47)$$

$$p_r = -\frac{\partial \mathcal{L}}{\partial \dot{r}} = \left(1 - \frac{2M}{r}\right)^{-1} \dot{r} \quad (8.48)$$

$$p_\theta = -\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = r^2 \dot{\theta} \quad (8.49)$$

$$p_\phi = -\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = r^2 \sin^2 \theta \dot{\phi}. \quad (8.50)$$

The resulting Hamiltonian is

$$\mathcal{H} = p_t \dot{t} - (p_r \dot{r} + p_\theta \dot{\theta} + p_\phi \dot{\phi}) - \mathcal{L} = \mathcal{L}. \quad (8.51)$$

The equality of the Hamiltonian and the Lagrangian signifies that there is no potential energy, as is evident from the definition of the energy integral. This fact indicates that both are constant

$$\mathcal{H} = \mathcal{L} = \text{const.} \quad (8.52)$$

By rescaling the affine parameter λ , we can arrange that $2\mathcal{C}$ has the value +1 for time-like geodesics, and zero for null geodesics (space-like geodesics are not covered).

Further integrals of motion follow from the equations

$$\frac{dp_t}{d\tau} = \frac{\partial \mathcal{L}}{\partial t} = 0 \quad (8.53)$$

$$\frac{dp_\phi}{d\tau} = \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (8.54)$$

Thus we find

$$p_t = \left(1 - \frac{2M}{r}\right) \frac{dt}{d\tau} = E = \text{const} \quad (8.55)$$

and

$$p_\phi = r^2 \sin^2 \theta \frac{d\phi}{d\tau} = L = \text{const.} \quad (8.56)$$

Moreover, from the equation of motion

$$\frac{dp_\theta}{d\tau} = \frac{d}{d\tau}(r^2 \dot{\theta}) = -\frac{\partial \mathcal{C}}{\partial \theta} = (r^2 \sin \theta \cos \theta) \left(\frac{d\phi}{d\tau}\right)^2, \quad (8.57)$$

it follows that $\dot{\theta} = 0$ when we assign the value $\pi/2$ to θ . Then $\ddot{\theta} = 0$, and θ will remain constant at the assigned value. We conclude that **the geodesic is described in invariant plane**. Then we find³

$$p_\phi = r^2 \frac{d\phi}{d\tau} = L = \text{const.} \quad (8.58)$$

³ The three constants of motion E , L and \mathcal{H} , which is essentially the normalization of the four-momenta, are a consequence of the existence of two Killing fields $K = \partial_t$ and $m = \partial_\phi$.

8.2.2 The Effective Potential for Equatorial Motion

With these conditions, the constancy of the Lagrangian gives

$$\frac{E^2}{1 - 2M/r} - \frac{\dot{r}^2}{1 - 2M/r} - \frac{L^2}{r^2} = 2\mathcal{L} = +1, 0, \quad (8.59)$$

depending on whether we are considering time-like or null geodesics. For time-like geodesics we find the two integrals of motion

$$\boxed{\left(\frac{dr}{d\tau}\right)^2 + \left(1 - \frac{2M}{r}\right)\left(1 + \frac{L^2}{r^2}\right) = E^2} \quad (8.60)$$

and

$$\boxed{\frac{d\phi}{d\tau} = \frac{L}{r^2}.} \quad (8.61)$$

The first equation is often written in terms of an effective potential in the form

$$\boxed{\left(\frac{dr}{d\tau}\right)^2 = E^2 - V^2,} \quad (8.62)$$

where

$$V^2 = \left(1 - \frac{2M}{r}\right)\left(1 + \frac{L^2}{r^2}\right). \quad (8.63)$$

This generalizes the Newtonian effective potential including the centrifugal term (Fig. 8.6).

Let us examine the kinds of possible orbits, as illustrated in Fig. 8.6. There are different curves $V(r)$ for different values of L/GM ; for any one of these curves, the behavior of the orbit can be judged by comparing the E^2 to $V^2(r)$. The general behavior of the particle will be to move in the potential until it reaches a “turning point” where $V(r) = E$; then it will begin moving in the other direction. Sometimes there may be no turning point to hit, in which case the particle just keeps going. In other cases the particle may simply move in a circular orbit at radius $r_c = \text{const}$; this can happen if the potential is flat, $dV/dr = 0$. Differentiating $V(r)$, we find that the circular orbits occur when

$$r_c^2 - \frac{L^2}{GM} r_c + 3\gamma L^2 = 0. \quad (8.64)$$

$\gamma = 0$ in Newtonian gravity and $\gamma = 1$ in general relativity. Circular orbits will be stable, if they correspond to a minimum of the potential, and unstable if they correspond to a maximum. Bound orbits which are not circular will oscillate around

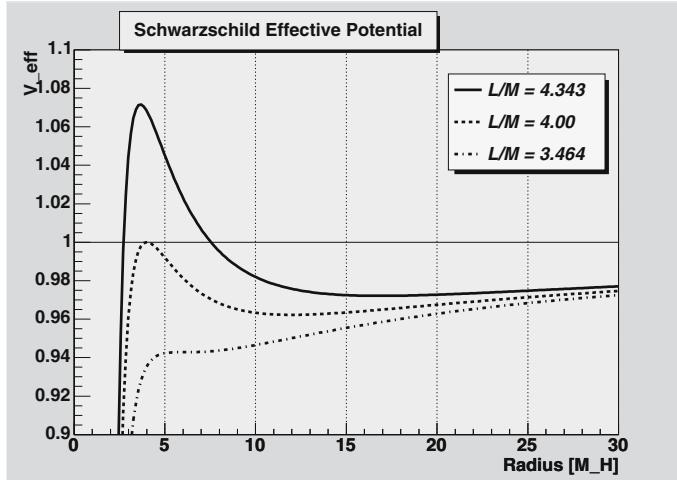


Fig. 8.6. Effective potential as a function of radius for various values of the angular momentum L/M . The abscissa is a dimensionless radius in units of gravitational radii, and GM/c is the natural unit for the specific angular momentum L of a particle. In distinction to the Newtonian effective potential, the relativistic effective potential attains a maximum for $L/M > 3.464$ and then vanishes at the Schwarzschild radius

the radius of the stable circular orbit. In Newtonian gravity, we find that circular orbits appear at

$$r_c = \frac{L^2}{GM}, \quad (8.65)$$

stating that at this radius the centrifugal force is balanced by gravity attraction.

In general relativity, the situation is different, but only for r sufficiently small. Since the difference resides in the term $-GML^2/r^3$, as $r \rightarrow \infty$, the behaviors are identical in the two theories. But as $r \rightarrow 0$, the potential goes to $-\infty$, rather than $+\infty$, as in the Newtonian case. At $r = 2GM/c^2$, the potential is always zero; inside this radius is the black hole. For massless particles, there is always a barrier (except for $L = 0$, for which the potential vanishes identically), but a sufficiently energetic photon will nevertheless go over the barrier and be dragged inexorably down to the center. At the top of the barrier there are unstable circular orbits.

For massive particles there are once again different regimes depending on the angular momentum. The circular orbits are at

$$r_c = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM}. \quad (8.66)$$

For large L there will be two circular orbits, one stable and one unstable

$$r_c \simeq \frac{L^2 \pm L^2(1 - 6G^2M^2/L^2)}{2GM} = \left(\frac{L^2}{GM}, 3GM \right). \quad (8.67)$$

In this limit, the stable circular orbit becomes farther and farther away, while the unstable one approaches $3GM$, a behavior which parallels the massless case. As we decrease L , the two circular orbits come closer together; they coincide when the discriminant vanishes

$$L = \sqrt{12} GM, \quad (8.68)$$

for which

$$r_c = r_{\text{ms}} = 6 GM/c^2 = 3R_S. \quad (8.69)$$

It disappears entirely for smaller L . Thus $6GM/c^2$ is the **smallest possible radius of a stable circular orbit in the Schwarzschild metric**. It is called the **radius of marginal stability, or ISCO (innermost stable circular orbit)**. There are also unbound orbits, which come in from infinity and turn around, and bound but noncircular ones, which oscillate around the stable circular radius. Note that such orbits, which would describe exact conic sections in Newtonian gravity, will not do so in GR, although we would have to solve the equation for d/dt to demonstrate it. Finally, there are orbits which come in from infinity and continue all the way in to $r = 0$; this can happen either if the energy is higher than the barrier, or for $L < \sqrt{12} GM$, when the barrier goes away entirely. We have therefore found that the

Schwarzschild solution possesses stable circular orbits for $r > 6GM/c^2$ and unstable circular orbits for $3GM/c^2 < r < 6GM/c^2$.

It is important to remember that these are only the geodesics; there is nothing to stop an accelerating particle from dipping below $r = 3GM/c^2$ and emerging, as long as it stays beyond $r = 2GM/c^2$.

8.2.3 Orbital Equation and Bound Orbits in Schwarzschild Spacetime

We now discuss exact solutions for the orbital motion in the equatorial plane. By considering r as a function of ϕ instead of τ , we obtain the equation

$$\left(\frac{dr}{d\phi} \right)^2 = (E^2 - 1) \frac{r^4}{L^2} + \frac{2M}{L^2} r^3 - r^2 + 2Mr. \quad (8.70)$$

We now introduce the variable $u \equiv 1/r$, as in the analysis of the Keplerian orbits in the Newtonian theory. With this replacement, we obtain the fundamental equation

$$\boxed{\left(\frac{du}{d\phi} \right)^2 = 2M u^3 - u^2 + \frac{2M}{L^2} u - \frac{1 - E^2}{L^2}.} \quad (8.71)$$

This equation determines the geometry of the geodesics in the invariant plane. Once it has been solved for $u = u(\phi)$, the solution can be completed by direct quadratures of the equations

$$\frac{d\tau}{d\phi} = \frac{1}{Lu^2} \quad (8.72)$$

$$\frac{dt}{d\phi} = \frac{E}{Lu^2(1 - 2Mu)}. \quad (8.73)$$

Bound Orbits

The solutions of (8.71) will depend on whether $E^2 < 1$ or $E^2 \geq 1$. This distinction is between bound orbits and unbound orbits. Bound orbits are governed by an equation

$$\left(\frac{du}{d\phi} \right)^2 = f(u), \quad (8.74)$$

where $f(u)$ is given by

$$f(u) = 2Mu^3 - u^2 + \frac{2M}{L^2}u - \frac{1 - E^2}{L^2}. \quad (8.75)$$

It is clear that the geometry of the geodesics will be determined by the positions of the roots $f(u) = 0$. Since $f(u)$ is cubic in u , there are two possibilities: either all roots are real, or one of them is real and the two remaining are complex-conjugate ones. Let u_1, u_2, u_3 denote the roots of $f(u) = 0$. Then we have

$$u_1 u_2 u_3 = (1 - E^2)/2ML^2 \quad (8.76)$$

and

$$u_1 + u_2 + u_3 = 1/2M. \quad (8.77)$$

Since $1 - E^2 > 0$, it must allow for one positive real root. From the further facts that $f < 0$ for $u = 0$ and $f(u) \rightarrow \pm\infty$ for $u \rightarrow \pm\infty$, we can distinguish between **five cases**:

1. Three different real roots $u_1 < u_2 < u_3$: There exist two distinct orbits confined to the interval $u_1 \leq u \leq u_2$ and $u > u_3$, i.e. an orbit which oscillates between two extreme values of r and an orbit, starting at a certain aphelion distance given by $1/u_3$ plunges into the singularity at $r = 0$, i.e. $u \rightarrow \infty$. These two classes of orbits are called **orbits of the first and second kinds**. The orbits of the first kind are the relativistic analogues of the Keplerian orbits. The orbits of the second kind have no Newtonian analogue. Orbit of both kinds are most conveniently parametrized by an eccentricity e and a latus rectum l , similar to Newtonian orbits.
2. $u_1 = u_2$ is a double root: In this case, the orbit of the first kind is a stable **circular orbit**, while the orbit of the second kind still plunges into the singularity.
3. $u_2 = u_3$ is a double root: In this case the orbit of the first kind starts at a certain aphelion distance $1/u_1$ and approaches the circle of radius $1/u_3$ asymptotically, by spiralling around it an infinite number of times. The orbit of the second kind is a continuation of the orbit of the first kind in that it spirals away from the same circle (towards the center).
4. $u_1 = u_2 = u_3$: All three roots coincide. This allows for an unstable circular orbit of radius $1/u_1$.
5. Only one real root u_1 : In this case we only have one class of orbits: they all plunge into the singularity after starting from a certain aphelion distance. They are similar to radial geodesics.

Orbits of the First Kind

In this case all three roots are positive, and we can write them as

$$u_1 = \frac{1}{l}(1 - e) \quad (8.78)$$

$$u_2 = \frac{1}{l}(1 + e) \quad (8.79)$$

$$u_3 = \frac{1}{2M} - \frac{2}{l}. \quad (8.80)$$

The semilatus rectum l is some positive constant⁴ and the eccentricity $e < 1$ for $u_1 > 0$, as required by the condition $E^2 < 1$.

The conformity with the ordering $u_1 < u_2 \leq u_3$ requires

$$\frac{1}{2M} - \frac{2}{l} \geq \frac{1+e}{l}, \quad (8.84)$$

or

$$l \geq 2M(3 + e). \quad (8.85)$$

Defining a parameter

$$\mu \equiv M/l \quad (8.86)$$

one finds the inequality

$$\mu \leq \frac{1}{2(3 + e)} \quad \text{or} \quad 1 - 6\mu - 2\mu e \geq 0. \quad (8.87)$$

In these parameters $f(u)$ is written as

$$f(u) = 2M \left(u - \frac{1-e}{l} \right) \left(u - \frac{1+e}{l} \right) \left(u - \frac{1}{2M} + \frac{2}{l} \right). \quad (8.88)$$

⁴ For a Keplerian ellipse, the semilatus rectum l is the distance l measured from a focus such that

$$\frac{1}{l} = \frac{1}{2} \left(\frac{1}{r_+} + \frac{1}{r_-} \right), \quad (8.81)$$

where $r_+ = a(1 + e)$ and $r_- = a(1 - e)$ are the aphelion and perihelion positions of the orbit. Plugging the values of r_+ and r_- into the equation for l gives

$$\frac{1}{l} = \frac{1}{a(1 - e^2)}, \quad (8.82)$$

and therefore

$$r_+ = \frac{l}{1-e}, \quad r_- = \frac{l}{1+e}. \quad (8.83)$$

This justifies the ansatz for the roots u_1 and u_2 .

This has to agree with the original form of the function, giving the relations

$$\frac{M}{L^2} = \frac{1}{l^2} [l - M(3 + e^2)] \quad (8.89)$$

$$\frac{1 - E^2}{L^2} = \frac{1}{l^3} [(l - 4M)(1 - e^2)] \quad (8.90)$$

or, in terms of μ

$$\frac{1}{L^2} = \frac{1}{lM} [1 - \mu(3 + e^2)] \quad (8.91)$$

$$\frac{1 - E^2}{L^2} = \frac{1}{l^2} (1 - 4\mu)(1 - e^2). \quad (8.92)$$

It follows from these equations that $\mu < 1/(3 + e^2)$ and $\mu < 1/4$.

As in the Keplerian problem, we now make the **ansatz**

$$u = \frac{1}{l} (1 + e \cos \chi).$$

(8.93)

χ is now a kind of **relativistic anomaly**. At aphelion, $\chi = \pi$, we find $u = (1 - e)/l$ and at perihelion, $\chi = 0$, $u = (1 + e)/l$. This substitution leads to the equation

$$\begin{aligned} \left(\frac{d\chi}{d\phi} \right)^2 &= 1 - 2\mu(3 + e \cos \chi) \\ &= (1 - 6\mu + 2\mu e) - 4\mu e \cos^2(\chi/2), \end{aligned} \quad (8.94)$$

or alternatively to

$$\pm \frac{d\chi}{d\phi} = \sqrt{1 - 6\mu + 2\mu e} \sqrt{1 - k^2 \cos^2(\chi/2)}, \quad (8.95)$$

where

$$k^2 = \frac{4\mu e}{1 - 6\mu + 2\mu e}. \quad (8.96)$$

The solution for ϕ can be expressed in terms of the Jacobian integral

$$F(\psi, k) = \int_0^\psi \frac{d\gamma}{\sqrt{1 - k^2 \sin^2 \gamma}}, \quad (8.97)$$

where

$$\psi = \frac{1}{2}(\pi - \chi). \quad (8.98)$$

Thus we may write

$$\boxed{\phi = \frac{2}{\sqrt{1 - 6\mu + 2\mu e}} F(\pi/2 - \chi/2, k),} \quad (8.99)$$

where the origin of ϕ has been chosen at aphelion passage where $\chi = \pi$. The perihelion passage occurs at $\chi = 0$, where $\psi = \pi/2$. A particular example is shown in Fig. 8.7 for the case $e = 0.5$ and $l = 10.0$.

The solution can be completed by the expressions for the proper time τ and the coordinate time t

$$\tau = \frac{1}{L} \int \frac{d\phi}{u^2} = \frac{1}{L} \int \frac{d\phi}{d\chi} \frac{d\chi}{u^2} \quad (8.100)$$

and

$$t = \frac{E}{L} \int \frac{d\phi}{d\chi} \frac{d\chi}{u^2(1 - 2Mu)}. \quad (8.101)$$

Post-Newtonian Corrections and Perihelion Advance

The first-order corrections to the Keplerian orbits of the Newtonian theory can readily be deduced from equation (8.94). Under normal conditions, the parameter $\mu = M/L$

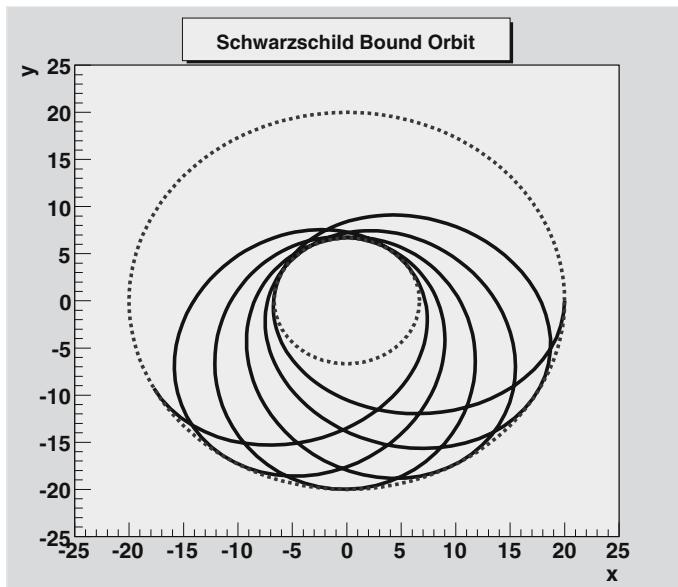


Fig. 8.7. Bound orbits of the first kind near a black hole, $E^2 < 1$. Shown are orbits with eccentricity $e = 0.5$ and $l = 11$

is a very small quantity – it is essentially the ratio of the gravitational radius M to the major axis of a planetary orbit or a binary star orbit. So we expand equation (8.94) to first order in μ to obtain

$$-d\phi = d\chi (1 + 3\mu + \mu e \cos \chi) \quad (8.102)$$

or in integrated form

$$-\phi = (1 + 3\mu)\chi + \mu e \sin \chi + \text{const.} \quad (8.103)$$

From this we infer that the change in ϕ after one complete revolution during which χ changes by 2π is $2\pi(1 + 3\mu)$. Therefore, the **advance of the perihelion**, $\Delta\phi$, per revolution is

$$\boxed{\Delta\phi = \frac{6\pi M}{l} = \frac{6\pi GM}{c^2 a(1 - e^2)}}.$$

(8.104)

a is the semimajor axis of the Keplerian ellipse. This is the standard result first derived by Einstein.

For the motion of Mercury around the Sun, the relevant orbital parameters are

$$\frac{GM_{\odot}}{c^2} = 1.48 \times 10^3 \text{ m}, \quad a = 5.79 \times 10^{10} \text{ m}, \quad e = 0.2056. \quad (8.105)$$

This gives

$$\Delta\phi_{\text{Mercury}} = 5.01 \times 10^{-7} \text{ radian/orbit} = 0.103 \text{ arcsec/orbit}. \quad (8.106)$$

It is customary to express this in terms of precession per century: Mercury orbits once every 88 days, yielding

$$\boxed{\Delta\phi_{\text{Mercury}} = 43.0 \text{ arcsec/century}.}$$

(8.107)

The major axis of Mercury's orbit precesses at a rate of 43.0 arcsecs every 100 years. The observed value is however 5601 arcsecs/100 years. Most of this is due to the precession of the equinoxes in our geocentric coordinate system, exactly 5025 arcsecs/100 years. The gravitational perturbations of the other planets contribute an additional 532 arcsecs/100 years, leaving 43 arcsecs/100 years to be explained by Einstein. Einstein was certainly very enthusiastic about these numbers when he figured this out for the first time. All the other planets show smaller perihelion precession due to the increasing semimajor axis.

8.3 The Kerr Black Hole

In a collapse situation, asymmetries of the matter distribution are rapidly radiated away by gravitational waves. One expects therefore that the collapse leads to a stationary and axisymmetric configuration, which is given by two Killing vectors. The final state of the ultimate collapse should be represented by a kind of ground state of Einstein's equations.

8.3.1 Kerr Black Hole in Boyer–Lindquist Coordinates

According to the considerations in Chap. 7, we may seek for a global vacuum solution with a stationary and axisymmetric line element of the form (see Sect. 7.1) given in quasispherical coordinates (t, ϕ, r, θ)

$$ds^2 = -\alpha^2 dt^2 + R^2 (d\phi - \omega dt)^2 + \exp(2\mu_2) dr^2 + \exp(2\mu_3) d\theta^2. \quad (8.108)$$

The five functions $v = \ln \alpha$, ω , $\psi = \ln R$, μ_2 and μ_3 only depend on the spatial coordinates r and θ . One has to remember that the two functions μ_2 and μ_3 are essentially not independent, but can be reduced to one by means of a gauge transformation.

In Boyer–Lindquist (BL) coordinates (t, ϕ, r, θ) the solution found by Kerr in 1963 can be expressed in the following form [86]

$$\alpha^2 \equiv e^{2v} = \frac{\varrho^2 \Delta}{\Sigma^2} \quad (8.109)$$

$$R \equiv e^\psi = \frac{\Sigma}{\varrho} \sin \theta \quad (8.110)$$

$$\omega = \frac{2aMr}{\Sigma^2} \quad (8.111)$$

$$e^{\mu_2} = \frac{\varrho}{\sqrt{\Delta}}, \quad e^{\mu_3} = \varrho. \quad (8.112)$$

It is a tradition to use the following polynomials of r and $\cos \theta$

$$\Delta = r^2 - 2Mr + a^2 \quad (8.113)$$

$$\varrho^2 = r^2 + a^2 \cos^2 \theta \quad (8.114)$$

$$\Sigma^2 = (r^2 + a^2)^2 - a^2 \Delta \sin^2 \theta. \quad (8.115)$$

This solution is uniquely given by **two parameters**: the mass M of the source and the Kerr parameter a , which is related to the angular momentum of the source, $J = aM$. In physical units, the mass is given in terms of the gravitational radius GM/c^2 , and similarly for the angular momentum, a is in units of GM/c^2 . This metric is asymptotically flat and approaches the Schwarzschild metric in the limit $a \rightarrow 0$.

8.3.2 A Short Derivation of the Kerr Solution

The original derivation of the solution by Kerr is quite implicit and cumbersome. Chandrasekhar [4] has given a more intuitive way towards a derivation, which is, however, still quite lengthy.

The Equations in BL Coordinates

As derived in Sect. 7.2, Einsteins's equations provide the following highly nonlinear equations for the five metric functions in the case of a global vacuum solution

$$\bar{\Delta}\alpha - \frac{R^2}{2\alpha} (\nabla\omega \cdot \nabla\omega) = 0 \quad (8.116)$$

$$\text{Div} \left(\frac{R}{\alpha} \nabla\omega \right) = 0 \quad (8.117)$$

$$\alpha \text{Div} \left(\frac{1}{\alpha} \nabla\psi \right) + \frac{R^2}{2\alpha^2} (\nabla\omega \cdot \nabla\omega) = 0 \quad (8.118)$$

$$\Delta(\mu_2, \mu_3) - \frac{R^2}{2\alpha^2} (\nabla\omega \cdot \nabla\omega) - \frac{2}{\alpha} \Psi_A (\nabla_A \alpha) = 0. \quad (8.119)$$

The last equation results from the Einstein tensors $R_{00} - R_{11} + R_{22} + R_{33}$. In addition, we also have to satisfy the constraint equation $R_{23} = 0$

$$\begin{aligned} R_{23} = & -\frac{1}{R} \nabla_3(R\Psi_2) + \Psi_3(\nabla_2\mu_3) + \frac{R^2}{2\alpha^2} (\nabla_3\omega)(\nabla_2\omega) \\ & -\frac{1}{\alpha} (\nabla_3\mu_2)(\nabla_2\alpha) - \frac{1}{\alpha} \nabla_2(\nabla_3\alpha) = 0. \end{aligned} \quad (8.120)$$

The first equation is the relativistic Poisson equation which shows that the extrinsic curvature $\nabla\omega$ acts as a source for the redshift factor α

$$\bar{\Delta}\alpha = \frac{R^2}{2\alpha} (\nabla\omega \cdot \nabla\omega). \quad (8.121)$$

The term on the right-hand side represents an effective energy density. The key equation is the second equation, which tells us that the vorticity of absolute space is conserved. It has no explicit source term. This equation can be seen in analogy to Ampère's equation, when expressed in terms of the vector potential $\Psi = RA_\phi$. In vacuum, this equation has the same form

$$\text{Div} \left(\frac{R}{\alpha} \nabla\Psi \right) = 0. \quad (8.122)$$

Similar to Maxwell's theory, the source of the frame-dragging potential is a ring current of orbiting mass elements of infinite density (a so-called ring singularity), as will be shown by analyzing the Riemann curvature tensors. This ring singularity generates the frame-dragging potential ω . Quantum effects are expected to smooth out the singularity. The third equation is of similar type, the axial curvature is also generated by frame-dragging. The last equation determines the meridional curvature, i.e. the two functions μ_2 and μ_3 in terms of frame-dragging and gravitational forces. It is therefore not astonishing that this nonlinear system of partial differential equations has an analytic solution.

For the first equation we need the Laplacian on the hypersurface with metric γ for $\alpha = \exp \nu$

$$\begin{aligned}\sqrt{\gamma} \bar{\Delta} \alpha &= \partial_A [\exp(\nu + \psi + \mu_2 + \mu_3) g^{AB} \partial_B \nu] \\ &= [\exp(\nu + \psi + \mu_3 - \mu_2) \nu_{,2}]_{,2} + [\exp(\nu + \psi + \mu_2 - \mu_3) \nu_{,3}]_{,3}.\end{aligned}\quad (8.123)$$

In total, the first equation is equivalent to

$$\begin{aligned}&[\exp(\nu + \psi + \mu_3 - \mu_2) \nu_{,2}]_{,2} + [\exp(\nu + \psi + \mu_2 - \mu_3) \nu_{,3}]_{,3} \\ &= \frac{1}{2} \exp(3\nu - \psi) [\exp(\mu_3 - \mu_2) (\omega_{,2})^2 + \exp(\mu_2 - \mu_3) (\omega_{,3})^2].\end{aligned}\quad (8.124)$$

The second equation determines the frame-dragging potential

$$[\exp(3\psi - \nu - \mu_2 + \mu_3) \omega_{,2}]_{,2} + [\exp(3\psi - \nu + \mu_2 - \mu_3) \omega_{,3}]_{,3} = 0. \quad (8.125)$$

For the following, it is useful to introduce the function

$$\beta = \psi + \nu. \quad (8.126)$$

The third equation determines the radius function

$$\begin{aligned}&[\exp(\beta + \mu_3 - \mu_2) \psi_{,2}]_{,2} + [\exp(\beta + \mu_2 - \mu_3) \psi_{,3}]_{,3} \\ &= -\frac{1}{2} \exp(3\psi - \nu) [\exp(\mu_3 - \mu_2) (\omega_{,2})^2 + \exp(\mu_2 - \mu_3) (\omega_{,3})^2].\end{aligned}\quad (8.127)$$

The sum and the difference between the first equation (8.124) and the third equation (8.128) give the equations for β and $\psi - \nu$

$$[\exp(\mu_3 - \mu_2) (\exp \beta)_{,2}]_{,2} + [\exp(\mu_2 - \mu_3) (\exp \beta)_{,3}]_{,3} = 0 \quad (8.128)$$

$$\begin{aligned}&[\exp(\beta + \mu_3 - \mu_2) (\psi - \nu)_{,2}]_{,2} + [\exp(\beta + \mu_2 - \mu_3) (\psi - \nu)_{,3}]_{,3} \\ &= -\exp(3\psi - \nu) [\exp(\mu_3 - \mu_2) (\omega_{,2})^2 + \exp(\mu_2 - \mu_3) (\omega_{,3})^2].\end{aligned}\quad (8.129)$$

The equation for $R_{23} = 0$ can explicitly be written as

$$\begin{aligned}(\nu + \psi)_{,2,3} - (\nu + \psi)_{,2} \mu_{2,3} - (\nu + \psi)_{,3} \mu_{3,2} + \psi_{,2} \psi_{,3} + \nu_{,2} \nu_{,3} \\ = \frac{1}{2} \exp(2\psi - 2\nu) \omega_{,2} \omega_{,3}.\end{aligned}\quad (8.130)$$

These five equations look quite desperate, but Chandrasekhar [4] has shown that the equations for the five functions can be split into two independent pairs of equations and one equation for $\mu_2 + \mu_3$. The essential assumption is the existence of a horizon which is in Boyer–Lindquist coordinates just given by a surface $r = \text{const}$. This means that BL coordinates are horizon-adapted coordinates. It is by no means clear that the horizon is not given by a general two-surface $\mathcal{N}(r, \theta) = \text{const}$ in any other quasispherical coordinate system.

(i) The Existence of a Horizon

In accordance with the assumed stationarity and axisymmetry, the equation for the event horizon is of the form

$$N(x^2, x^3) = 0 \quad (8.131)$$

with the condition being a null surface

$$g^{AB} N_{,A} N_{,B} = 0, \quad A, B = 2, 3. \quad (8.132)$$

This implies for the chosen metric

$$e^{2(\mu_3 - \mu_2)} N_{,r}^2 + N_{,\theta}^2 = 0 \quad (8.133)$$

for quasispherical coordinates $x^2 = r$ und $x^3 = \theta$. If the horizon is given by a surface $r = \text{const}$, $N_{,\theta}$ vanishes on the horizon. We can now use the gauge freedom to specify a relation between μ_2 and μ_3 as a mere function of radius

$$e^{2(\mu_3 - \mu_2)} \equiv \Delta(r), \quad (8.134)$$

where $\Delta(r)$ is some function to be found. The equation for the null surface at position $r = r_H$ requires therefore

$$\Delta(r_H) = 0. \quad (8.135)$$

The equation for β can then suitably be solved. Since Δ is at our disposal, β can be solved independently of the other functions. The central problem is to solve the equations for $\psi - v$ and ω . The solution for $\mu_2 + \mu_3$ presents no difficulty, once the solutions for $\psi - v$ and ω are known.

For the following, we introduce the functions

$$\beta \equiv \psi + v, \quad \chi \equiv e^{v-\psi}. \quad (8.136)$$

The second condition that this null surface is spanned by the Killing vectors ∂_t and ∂_ϕ requires that the determinant of this subspace also vanishes at the horizon

$$e^{2\beta}(r_H) = 0. \quad (8.137)$$

For this reason we can try an ansatz of the form

$$e^\beta = \sqrt{\Delta} f(r, \theta) \quad (8.138)$$

with a function f regular on the surface $\Delta = 0$ and on the axis $\theta = 0$. In fact, we can even try a separable ansatz with $f = f(\theta)$ and demonstrate that this leads to solutions of Einstein's equations. It is one of the mysteries of the Kerr geometry that many equations are separable (see, e.g. the Hamilton–Jacobi equation for geodesics).

With the above ansatz for β , the equation for β is simply

$$\left[\sqrt{\Delta} (e^\beta)_{,2} \right]_{,2} + \left[\frac{1}{\sqrt{\Delta}} (e^\beta)_{,3} \right]_{,3} = 0. \quad (8.139)$$

This leads to

$$\left[\sqrt{\Delta} (\sqrt{\Delta})_{,r} \right]_{,r} + \frac{1}{f} f_{,\theta\theta} = 0. \quad (8.140)$$

Since f must be regular on the axis and we want to have the usual expression for the cylindrical radius in the asymptotic region, the only choice is

$$f(\theta) = \sin \theta \quad (8.141)$$

and therefore Δ satisfies a very simple equation

$$\Delta_{,rr} = 2. \quad (8.142)$$

The general solution contains two free parameters M and a

$$\boxed{\Delta(r) = r^2 - 2Mr + a^2} \quad (8.143)$$

which are chosen in accordance with its later interpretation as mass M and angular momentum parameter a . As it turns out, these are the only free parameters of the solution, since the other equations are highly nonlinear and do not allow the introduction of further free parameters. Both parameters M and a have the dimension of a length, the corresponding mass is denoted as M_H with the relation $M = GM_H/c^2$. Thus, with this choice of the gauge, we have determined two of the five functions

$$e^{\mu_3 - \mu_2} = \sqrt{\Delta} \quad (8.144)$$

$$e^\beta = e^{\psi + \nu} = \sqrt{\Delta} \sin \theta. \quad (8.145)$$

(ii) Three Key Equations

Since two of the five functions are determined by the equations (8.144) and (8.145), we are left with three equations for ω , α and ψ

$$\left[\exp(4\psi) \frac{\omega_{,r}}{\sin \theta} \right]_{,r} + \left[\exp(4\psi) \frac{\omega_{,\theta}}{\Delta \sin \theta} \right]_{,\theta} = 0 \quad (8.146)$$

$$[(\Delta \sin \theta) \nu_{,r}]_{,r} + [\sin \theta \nu_{,\theta}]_{,\theta} = \frac{\exp(4\psi)}{2\Delta \sin \theta} [\Delta(\omega_{,r})^2 + (\omega_{,\theta})^2] \quad (8.147)$$

$$[(\Delta \sin \theta) \psi_{,r}]_{,r} + [\sin \theta \psi_{,\theta}]_{,\theta} = -\frac{\exp(4\psi)}{2\Delta \sin \theta} [\Delta(\omega_{,r})^2 + (\omega_{,\theta})^2]. \quad (8.148)$$

These are highly nonlinearly coupled partial differential equations of second order, which, however, surprisingly can be solved by an ansatz with rational functions (this is another mystery of the Kerr solution)

$$\omega(r, \theta) = \frac{2aMr}{\Sigma^2} \quad (8.149)$$

$$R^2(r, \theta) = \exp(2\psi) = \frac{\Sigma^2 \sin^2 \theta}{\varrho^2} \quad (8.150)$$

$$\alpha^2(r, \theta) = \exp(2\nu) = \frac{\varrho^2 \Delta}{\Sigma^2} \quad (8.151)$$

$$\chi(r, \theta) = \exp(\nu - \psi) = \frac{\varrho^2 \sqrt{\Delta}}{\Sigma^2 \sin \theta}. \quad (8.152)$$

where Σ^2 and ϱ^2 are polynomials of r and $\cos \theta$, specified in the beginning. It can be shown that this ansatz in fact solves all three equations.

(iii) The Meridional Curvature

The last equation we have to solve is for the surface element $\exp(\mu_2 + \mu_3)$ of the meridional plane. Since $\exp \mu_2 = \exp \mu_3 / \sqrt{\Delta}$, the last equation provides us a second-order differential equation for $\exp \mu_3$. The solutions for the two meridional metric functions are therefore

$$\boxed{\exp(2\mu_2) = \varrho^2 / \Delta, \quad \exp(2\mu_3) = \varrho^2.} \quad (8.153)$$

With this last relation we have determined all five metric functions.

8.3.3 The Weyl–Papapetrou Form of the Kerr Metric

It is interesting that the metric can be expressed in terms of the functions Δ and χ

$$ds^2 = -\sqrt{\Delta} \sin \theta \left[\chi(dt)^2 - \frac{1}{\chi}(d\phi - \omega dt)^2 \right] + \frac{\exp(\mu_2 + \mu_3)}{\sqrt{\Delta}} [(dr)^2 + \Delta(d\theta)^2], \quad (8.154)$$

with the introduction of

$$\chi \equiv \exp(\nu - \psi). \quad (8.155)$$

This is also equivalent to

$$ds^2 = -\exp \beta \left[\chi(dt)^2 - \frac{1}{\chi}(d\phi - \omega dt)^2 \right] + \exp(2\mu) [(dr)^2 + \Delta(d\theta)^2] \quad (8.156)$$

Since β satisfies the simple Laplace equation

$$\nabla^2 \exp \beta = 0, \quad (8.157)$$

we may consider $\exp \beta$ as one of the coordinates and seek a coordinate transformation

$$(x^2, x^3) \rightarrow (\varrho, z) \quad (8.158)$$

such that

$$\exp(2\mu) [(dx^2)^2 + (dx^3)^2] \rightarrow f(\varrho, z) [(d\varrho)^2 + (dz)^2]. \quad (8.159)$$

Such a transformation is possible, if

$$(\varrho_{,2})^2 + (z_{,2})^2 = (\varrho_{,3})^2 + (z_{,3})^2 \quad (8.160)$$

and

$$\varrho_{,2}\varrho_{,3} + z_{,2}z_{,3} = 0. \quad (8.161)$$

These conditions can be satisfied by means of

$$\varrho_{,2} = z_{,3}, \quad \varrho_{,3} = -z_{,2}. \quad (8.162)$$

Since $\nabla^2 \varrho = 0$ for $\varrho \equiv \exp \beta$, these relations can be satisfied. With this transformation we have achieved then the **Weyl–Papapetrou form** of the Kerr metric with the explicit expression for the cylindrical radius

$$\varrho = \sqrt{\Delta} \sin \theta. \quad (8.163)$$

This shows that the cylindrical radius ϱ is only a suitable coordinate outside the horizon, where Δ vanishes.

8.3.4 Uniqueness of the Kerr Solution

The Kerr solution, which endows a stationary, axisymmetric, asymptotically flat spacetime with a smooth convex event horizon is just characterized by two parameters – the mass M and the angular momentum $J = aM$. The uniqueness of the Kerr metric for the description of black holes follows from a **theorem by Robinson** [345]:

Stationary axisymmetric solutions of Einstein's equations for the vacuum which satisfy

1. *are asymptotically flat,*
2. *contain a smooth convex horizon,*
3. *are nonsingular outside the horizon*

are uniquely specified by two parameters: the mass M and angular momentum J_H with $J_H < M_H^2$, and only by these two parameters.⁵

⁵ For the proof, see, e.g. [4], p. 292.

8.3.5 Global Properties of the Kerr Metric

The Kerr metric has various interesting properties which are shortly discussed.

Asymptotic Expansion

For vanishing Kerr parameter, $a = 0$, the Kerr metric becomes the static, spherically symmetric Schwarzschild metric. The leading order terms in asymptotic expansion are given by

$$\alpha^2 \simeq 1 - \frac{2M}{r}, \quad R \simeq r \sin \theta \quad (8.164)$$

$$\omega \simeq \frac{2aM}{r^3}, \quad e^{-2\mu_2} \simeq 1 - \frac{2M}{r}, \quad e^{\mu_3} \simeq r. \quad (8.165)$$

Therefore, the metric is asymptotically flat and the parameter M can be identified with the mass of the black hole. The total angular momentum follows from the specific angular momentum a , according to $J_H = aM_H$. Hence, the function ω can be interpreted as a potential that is unambiguously determined by the angular momentum of the rotating black hole. As we have seen in the previous chapter, there exists no analogue in Newtonian gravity. The quantity $2\pi R = 2\pi\sqrt{g_{\phi\phi}}$ represents the circumference of the cylinders that are concentric to the axis of symmetry, the axis of rotation of the black hole located at $\theta = 0$.

Event Horizons

The horizon is a 2D surface of spherical topology, where the redshift factor vanishes (i.e. the redshift observed at infinity is infinite, $z \rightarrow \infty$)

$$\alpha_H = 0 \mapsto \Delta(r_\pm) = 0. \quad (8.166)$$

The equation $\Delta(r) = 0$ has in general two solutions. These are the **outer horizon** or **event horizon** of the Kerr black hole given by its radius

$$r_+ = M + \sqrt{M^2 - a^2}, \quad (8.167)$$

and the **inner horizon** or **Cauchy horizon**

$$r_- = M - \sqrt{M^2 - a^2}. \quad (8.168)$$

In physical units, one has to replace M by GM_H/c^2 and a by $a = a_*(GM_H/c^2)$. The total angular momentum of the black hole will be given by $J_H = M_H a$. Physically meaningful solutions have therefore to satisfy the condition $|a| \leq M$, or $|a_*| \leq 1$, otherwise no horizon will exist. The special case $a = M$ is known as the **extreme Kerr solution**. From this we get the maximal specific angular momentum of a Kerr black hole $j_{\max} = GM_H/c$, und $J_H \leq M_H j_{\max}$.

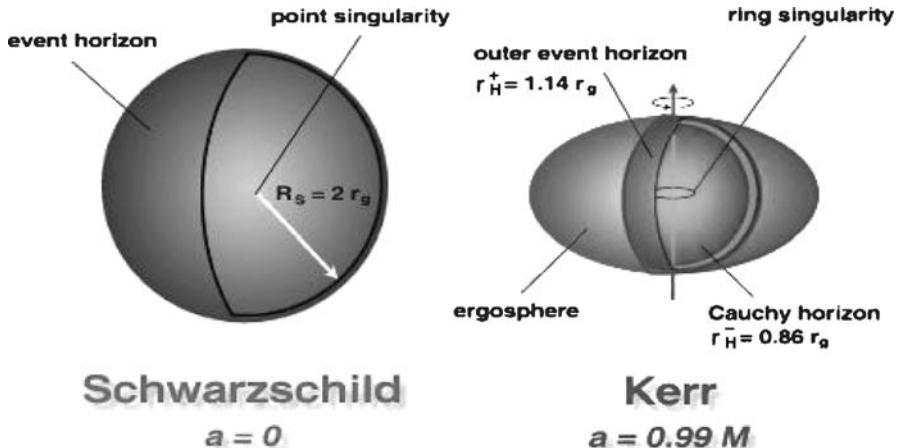


Fig. 8.8. Direct comparison of the structures of a Schwarzschild black hole (*left*) and an extreme Kerr black hole (*right*). Nonrotating black holes exhibit a point singularity. Rotating black holes have a ring singularity and an oblate ergosphere. Figure provided by A. Müller [304]

In consequence, there will be two horizons in general. The horizons are *null surfaces* because they are light-like. The gravitational redshift suppresses *any* emission at the event horizon. This results in the blackness of the black hole. Cauchy horizons are semipermeable surfaces: once trapped behind the Cauchy horizon there is no turning back. An observer crossing the Cauchy horizon witnesses the complete history of the outer world only in one instant of time. This is because the inner horizon represents a region of infinite blueshift (see, e.g. the time-like trajectories in the Penrose diagram 8.12).

The hypersurfaces with $r = r_{\pm}$ are **Killing horizons** of the Killing vector field

$$\boxed{\xi = k + \Omega_H m \quad , \quad \Omega_H = \frac{a}{r_{\pm}^2 + a^2}} . \quad (8.169)$$

At these surfaces we measure **surface gravities**

$$\boxed{\kappa_{\pm} = \frac{r_{\pm} - r_{\mp}}{2(r_{\pm}^2 + a^2)}} . \quad (8.170)$$

Intermezzo about Killing Horizons

A null hypersurface \mathcal{N} is called Killing horizon of a Killing vector field ξ , if ξ is normal to \mathcal{N} on \mathcal{N} . Let l be a normal to \mathcal{N} such that $l \cdot Dl = 0$ in an affine parametrization. Then, since $\xi = f l$ on \mathcal{N} , for some function f , it follows that

$$(\xi \cdot D)\xi = \kappa \xi \quad \text{on } \mathcal{N} , \quad (8.171)$$

where $\kappa = (i_\xi \cdot d) \ln f$ is called **surface gravity**. Since ξ is normal to \mathcal{N} , the Frobenius theorem implies

$$\xi_{[\mu} D_{\nu]} \xi_{\rho}]|_{\mathcal{N}} = 0. \quad (8.172)$$

For a Killing vector field ξ , $D_\mu \xi_\nu = D_{[\mu} \xi_{\nu]}$, i.e. the symmetric part vanishes. In this case, the above equation can be written as

$$\xi_\rho D_\mu \xi_\nu + (\xi_\mu D_\nu \xi_\rho - \xi_\nu D_\mu \xi_\rho) = 0. \quad (8.173)$$

Multiplying by $D^\mu \xi^\nu$, we get

$$\xi_\rho (D^\mu \xi^\nu) D_\mu \xi_\nu = -2(D^\mu \xi^\nu) \xi_\mu D_\nu \xi_\rho \quad (8.174)$$

or

$$\begin{aligned} \xi_\rho (D^\mu \xi^\nu) D_\mu \xi_\nu &= -2(\xi \cdot D \xi^\nu) D_\nu \xi_\rho \\ &= -2\kappa \xi \cdot D \xi_\rho \\ &= -2\kappa^2 \xi_\rho|_{\mathcal{N}}. \end{aligned} \quad (8.175)$$

Hence, except for points at which $\xi = 0$,

$$\boxed{\kappa^2 = -\frac{1}{2}(D^\mu \xi^\nu)(D_\mu \xi_\nu)|_{\mathcal{N}}} \quad (8.176)$$

One can then prove that κ is constant on Killing horizons, i.e.

$$(\xi \cdot \partial) \kappa^2 = -(D^\mu \xi^\nu) R_{\nu\mu\rho\sigma} \xi^\rho \xi^\sigma = 0 \quad (8.177)$$

due to the antisymmetry of the Riemann tensor. In a Kruskal spacetime, the surface gravity is given by

$$|\kappa| = \frac{c^3}{4GM}, \quad (8.178)$$

and the Killing field $\xi = k$ spans a Killing horizon.

In general there is no natural normalization of the surface gravity, since $\xi^2 = 0$. In an asymptotically flat spacetime, the requirement that $k^2 \rightarrow -1$ in the asymptotic region provides however a suitable normalization.

In order to see that the hypersurfaces \mathcal{N}_\pm in the Kerr geometry defined by $r = r_\pm$ are Killing horizons, one has to transform the Kerr metric into Kerr coordinates (v, χ, r, θ) defined by

$$dv = dt + \frac{r^2 - a^2}{\Delta} dr \quad (8.179)$$

$$d\chi = d\phi + \frac{a}{\Delta} dr, \quad (8.180)$$

giving the metric

$$\begin{aligned} ds^2 = & - \frac{\Delta - a^2 \sin^2 \theta}{\varrho^2} dv^2 + 2dv dr \\ & - \frac{2a \sin^2 \theta (r^2 + a^2 - \Delta)}{\varrho^2} dv d\chi - 2a \sin^2 \theta d\chi dr \\ & + \frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta}{\varrho^2} \sin^2 \theta d\chi^2 + \varrho^2 d\theta^2. \end{aligned} \quad (8.181)$$

The normals l_{\pm} to the hypersurfaces \mathcal{N}_{\pm} are found to be given by

$$l_{\pm} \propto \frac{r_{\pm}^2 + a^2}{r_{\pm}^2 + a^2 \cos^2 \theta} f_{\pm} \left(\frac{\partial}{\partial v} + \frac{a}{r_{\pm}^2 + a^2} \frac{\partial}{\partial \chi} \right). \quad (8.182)$$

First of all one can show that $l_{\pm}^2 = 0$, so that \mathcal{N}_{\pm} are null hypersurfaces. Then we define $\xi_{\pm} = \partial_v + a/(r_{\pm}^2 + a^2) \partial_{\chi}$ and \mathcal{N}_{\pm} turn out to be Killing horizons of ξ_{\pm} . A straightforward calculation gives then values of κ_{\pm} in equation (8.170).

Angular Velocity of the Horizon

The event horizon is a Killing horizon for the Killing field

$$\boxed{\xi = k + \Omega_H m}, \quad (8.183)$$

with $\xi^2 = 0$, where

$$\boxed{\Omega_H = \frac{a}{r_+^2 + a^2} = \frac{1}{2} \frac{a}{M} \frac{1}{r_+}} \quad (8.184)$$

is called the angular velocity of the horizon. This angular velocity is independent of the latitude, i.e. the horizon is rigidly rotating. Unfortunately, no orbiting spots can be observed on the horizon, since all radiation from the horizon suffers infinite redshift.

Ergosphere and Frame-Dragging

There is an intriguing difference between the Schwarzschild and the Kerr solution: in the Kerr spacetime there is no globally static observer. In contrast to Schwarzschild, in Kerr there is no globally time-like nonrotating Killing field.

Although k is time-like at infinity, it need not be time-like everywhere outside the horizon. For Kerr we find

$$k^2 = g_{tt} = - \frac{\Delta - a^2 \sin^2 \theta}{\varrho^2} = - \left(1 - \frac{2Mr}{r^2 + a^2 \cos^2 \theta} \right). \quad (8.185)$$

k is therefore time-like, provided

$$r^2 + a^2 \cos^2 \theta - 2Mr > 0. \quad (8.186)$$

For $M^2 > a^2$, this implies

$r > r_E(\theta) = M + \sqrt{M^2 - a^2 \cos^2 \theta}.$

(8.187)

The boundary of this region, $r = r_E(\theta)$, marks the transition of the coordinate t from a time-like to a space-like coordinate. This is called the *static limit* or **ergosphere**. The region $r_+ \leq r \leq r_E$ is the so-called **ergoregion** (Fig. 8.9).

The ergosphere intersects the horizon at the two poles, but lies outside the horizon for all other latitudes.

As can immediately be investigated from the formula for the static limit, the ergosphere has an angle dependence. Therefore, it has an oblate structure: in the equatorial plane the ergosphere starts always at the Schwarzschild radius, $2GM/c^2$; but it ends at the poles at the outer horizon, $r = r_+$. Stationarity means illustratively that hypersurfaces with $t = \text{const}$ can be shifted identically. But the corresponding Killing field is not globally time-like. In the vicinity of the hole it becomes space-like.

Within the ergosphere, the so-called ergoregion, everything must rotate! This includes observers, photons and magnetic field lines. The name of this phenomenon is the **frame-dragging effect**. Let us consider a velocity field of an observer (or a plasma particle) of the form

$$U = U^t(k + \Omega m), \quad \Omega = \frac{U^\phi}{U^t}. \quad (8.188)$$

Kerr
 $0 < a < 0.7$

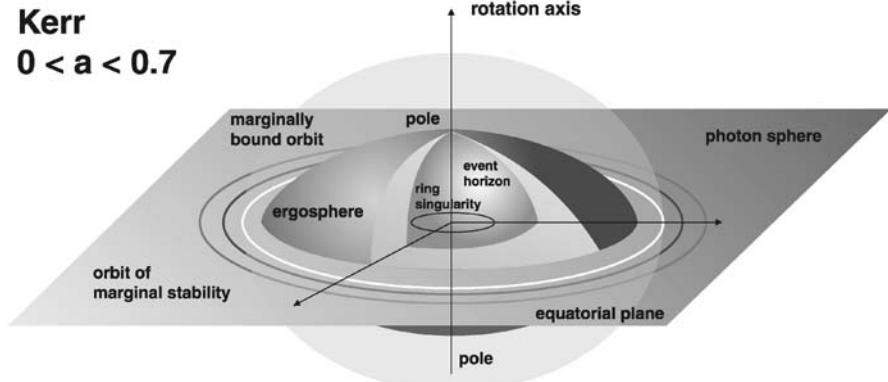


Fig. 8.9. Ergosphere and ergoregion of a Kerr black hole. Also shown is the orbit of marginal stability. This is the innermost stable circular orbit (ISCO) where stable rotation is possible. For smaller radii, the orbiter must fall into the hole or escape. When $a_* > 0.7$, the photon orbit moves inside the ergosphere (see Fig. 8.11). Figure provided by A. Müller [304]

This velocity field must be globally time-like which implies (in the signature $- + ++$)

$$g_{tt} + 2\Omega g_{t\phi} + \Omega^2 g_{\phi\phi} < 0, \quad (8.189)$$

and hence, Ω has to satisfy the limits (Fig. 8.10)

$$\Omega_- \leq \Omega \leq \Omega_+, \quad \Omega_- = \omega - \frac{c\alpha}{R}, \quad \Omega_+ = \omega + \frac{c\alpha}{R}.$$

(8.190)

Since $\alpha(r_+) = 0$, any observer (or plasma particle) rotates at the outer horizon with the angular velocity of the Kerr hole, Ω_H ,

$$\Omega_H \equiv \omega(r_+) = \frac{1}{2} \frac{a}{M} \frac{1}{r_+} = \frac{a}{r_+^2 + a^2}.$$

(8.191)

In the Kerr space we find special observers, called ZAMO (*zero angular momentum observer*) with angular velocity $\Omega = \omega$ (angular velocity with respect to fixed stars) which will have vanishing specific angular momentum

$$\begin{aligned} U_\phi &= g_{\phi\alpha} U^\alpha = g_{\phi\phi} U^\phi + g_{t\phi} U^t \\ &= U^t (\omega g_{\phi\phi} + g_{t\phi}) = U^t g_{\phi\phi} (\omega + \frac{g_{t\phi}}{g_{\phi\phi}}) = 0. \end{aligned} \quad (8.192)$$

It is therefore suitable to express physical observables with respect to this special observer system, also called Bardeen observers.

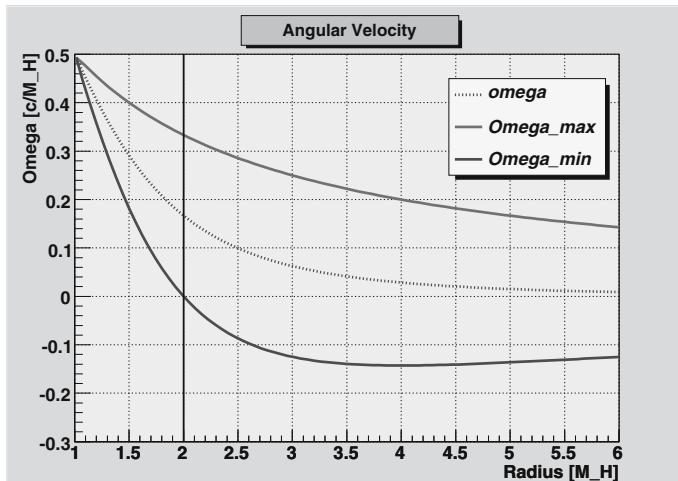


Fig. 8.10. Illustration of frame-dragging for a Kerr parameter, $a = 1.0$, in the equatorial plane. The *upper solid curve*, Ω_+ , represents the prograde limit of angular velocity; the *lower solid curve* limits the retrograde case, Ω_- . At the outer horizon, $r_+ = 1.0 GM/c^2$, the limits equal the frame-dragging potential ω (*dotted line*). Within one Schwarzschild radius (*ergoregion*), any form of matter must rotate with a fraction of the horizon's angular velocity

Characteristic Radii

The Kerr geometry exhibits a number of characteristic radii. Some of them have already been introduced, e.g. the inner and outer horizon, r_{\pm} , and ergosphere r_E .

Besides these radii, there are several others that are of special interest for accretion theory of black holes (Fig. 8.9). The radius of marginal stability will be derived in parametrized form (see Sect. 8.3)

$$r_{\text{ms}} = M \left(3 + Z_2 \mp \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)} \right) \quad (8.193)$$

$$Z_1 = 1 + \left(1 - \frac{a^2}{M^2} \right)^{1/3} \left(\left(1 + \frac{a}{M} \right)^{1/3} + \left(1 - \frac{a}{M} \right)^{1/3} \right) \quad (8.194)$$

$$Z_2 = \sqrt{3 \frac{a^2}{M^2} + Z_1^2}. \quad (8.195)$$

The solution of this equation determines the critical radius where stable rotation on circular orbits is possible. For $r < r_{\text{ms}}$, there are no stable orbits and the orbiting object is forced to fall into the hole or to escape. Two interesting limits are $r_{\text{ms}}(a = 1) = M$ (extreme Kerr) and $r_{\text{ms}}(a = 0) = 6M$ (the Schwarzschild case).

The marginally bound orbit, r_{mb} , is the characteristic radius where a test particle starts (as viewed from infinity) to be gravitationally bound by the black hole

$$r_{\text{mb}} = \left(\sqrt{M} + \sqrt{M - a} \right)^2. \quad (8.196)$$

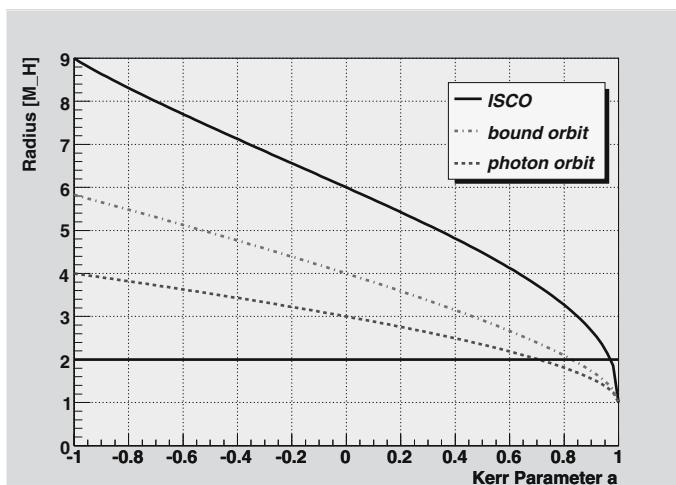


Fig. 8.11. Characteristic radii of a Kerr black hole in the equatorial plane as a function of the angular momentum a^* . Black holes are confined to the range $-1 \leq a^* \leq +1$. The *upper curve* represents the radius of marginal stable orbits, the *middle curve* the marginally bound orbit and the *lower curve* the photon orbit. The *solid horizontal line* gives the position of the ergosphere

Another radius marks the **photon sphere**,

$$r_{\text{ph}} = 2M \left[1 + \cos \left(\frac{2}{3} \arccos(-a/M) \right) \right]. \quad (8.197)$$

Photons may circulate around the Kerr black hole in this sphere, but the trajectories are unstable. All characteristic radii are shown in comparison and their dependence on the Kerr parameter a in Fig. 8.11. It is remarkable that nearly all these characteristic radii coincide for an extreme Kerr black hole, $a_* = +1$.

On the Ring Singularity

As in Schwarzschild, the horizons are no true singularities in Kerr. In order to find the true singularities, we have to calculate the components of the curvature tensor. This is left as an exercise to the reader.

As indicated in the discussion of the components of the Ricci tensor there exists a **mass current**. Due to axisymmetry of the Kerr metric this current is located in the equatorial plane. The derivation of this feature follows from the curvature tensor of the Kerr geometry. It tells us that the condition for the intrinsic singularity, the curvature singularity, is $\varrho(r, \theta) = 0$. With $\varrho^2 = r^2 + a^2 \cos^2 \theta$ one finds that $r = \cos \theta = 0$ must be simultaneously fulfilled. This can occur when $r = 0$ and $\theta = \pi/2$. Surprisingly, this can be interpreted as a ring singularity lying in the equatorial plane. **This intrinsic singularity is the source of the axisymmetric gravitational field of a rotating black hole.** Figure 8.12 sketches the two horizons that are separated by the ring singularity.

8.3.6 On the Conformal Structure of the Kerr Solution

Let us think about the structure of the full Kerr solution (for the basics of Penrose diagrams, see Sect. 8.1.4). Singularities seem to appear at both $\Delta = 0$ and $\varrho = 0$; let us turn our attention first to $\Delta = 0$. As in the Reissner–Nordström solution, there are three possibilities: (i) $M^2 > a^2$, (ii) $M^2 = a^2$, and (iii) $M^2 < a^2$. The last case features a naked singularity, and the extremal case $M^2 = a^2$ is unstable. Since these cases are of less physical interest, we will concentrate on $M^2 > a^2$. Then there are two radii at which Δ vanishes, given by $r = r_{\pm}$. Both radii are null surfaces which will turn out to be event horizons. The analysis of these surfaces proceeds in close analogy with the Reissner–Nordström case; it is straightforward to find coordinates which extend through the horizons.

Before rushing to draw Penrose diagrams, we need to understand the nature of the true curvature singularity; this does not occur at $r = 0$ in this spacetime, but rather at $\varrho = 0$. Since $\varrho^2 = r^2 + a^2 \cos^2 \theta$ is the sum of two manifestly nonnegative quantities, it can only vanish when both quantities are zero, or for $r = 0$ and $\theta = \pi/2$. This seems like a funny result, but remember that $r = 0$ is not a point in space, but a disk; the set of points $r = 0, \theta = \pi/2$ is actually the **ring at the edge of this**

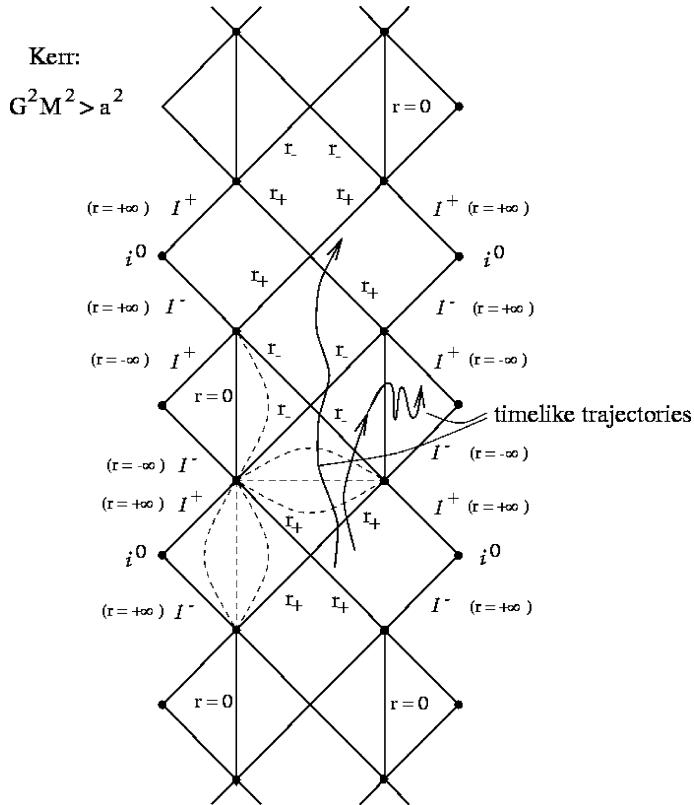


Fig. 8.12. The conformal structure (Penrose diagramm) for rotating black holes. This is quite similar to the Penrose diagramm for the Reissner–Nordström solution, except that one can move beyond the singularity

disk. The rotation has “softened” the Schwarzschild singularity, spreading it out over a ring.

What happens if you go inside the ring? A careful analytic continuation (which we will not perform) would reveal that you exit to another asymptotically flat spacetime, but not an identical copy of the one you came from. The new spacetime is described by the Kerr metric with $r < 0$. As a result, Δ never vanishes and there are no horizons. The Penrose diagram (Fig. 8.12) is much like that for Reissner–Nordström, except now you can pass through the singularity.

8.3.7 Ernst’s Equations for the Kerr Geometry

As we have seen in the derivation of the Kerr solution, there are two essential equations for the two functions χ and ω which can be written in terms of $\Delta(r)$ and $\delta(\theta) = \sin^2 \theta$, with $\chi \equiv \exp(v - \Psi)$,

$$\left(\frac{\Delta}{\chi}\chi_{,2}\right)_{,2} + \left(\frac{\delta}{\chi}\chi_{,3}\right)_{,3} = \frac{1}{\chi^2} [\Delta(\omega_{,2})^2 + \delta(\omega_{,3})^2] \quad (8.198)$$

$$\left(\frac{\Delta}{\chi^2}\omega_{,2}\right)_{,2} + \left(\frac{\delta}{\chi^2}\omega_{,3}\right)_{,3} = 0. \quad (8.199)$$

Obviously, there is a certain symmetry in these equations. The second equation can be solved in terms of a potential Φ defined as

$$\Phi_{,2} = \frac{\delta}{\chi} \omega_{,3}, \quad \Phi_{,3} = -\frac{\Delta}{\chi} \omega_{,2} \quad (8.200)$$

resulting in the Laplace-type equation

$$\left(\frac{\chi^2}{\delta}\Phi_{,2}\right)_{,2} + \left(\frac{\chi^2}{\Delta}\Phi_{,3}\right)_{,3} = 0. \quad (8.201)$$

In terms of this potential, the right-hand side of the first equation can be rewritten as

$$[\Delta(\log \chi)_{,2}]_{,2} + [\delta(\log \chi)_{,3}]_{,3} = \frac{\chi^2}{\Delta} (\Phi_{,2})^2 + \frac{\chi^2}{\delta} (\Phi_{,3})^2. \quad (8.202)$$

Introducing a function $\Psi \equiv \sqrt{\Delta\delta}/\chi$ one ends up with the two equations

$$\Psi [(\Delta\Psi_{,2} + (\delta\Psi_{,3})_{,3}] = \Delta [(\Psi_{,2})^2 - (\Phi_{,2})^2] + \delta [(\Psi_{,3})^2 - (\Phi_{,3})^2] \quad (8.203)$$

$$\Psi [(\Delta\Psi_{,2})_{,2} + (\delta\Psi_{,3})_{,3}] = 2\Delta\Psi_{,2}\Phi_{,2} + 2\delta\Psi_{,3}\Phi_{,3}. \quad (8.204)$$

These two equations can be combined then into one complex equation by introducing a **complex potential** $\mathcal{E} \equiv \Psi + i\Phi$ (Ernst [145], see also Sect. 7.6)

$$Re(\mathcal{E}) [(\Delta\mathcal{E}_{,2})_{,2} + (\delta\mathcal{E}_{,3})_{,3}] = \Delta(\mathcal{E}_{,2})^2 + \delta(\mathcal{E}_{,3})^2. \quad (8.205)$$

As a result, two of Einstein's equations are merged into one single complex potential equation. This equation has many interesting properties. So, for example, if \mathcal{E} is a solution, then \mathcal{E}^{-1} is also a solution. In addition, the combination $\mathcal{E}/(1 + ic\mathcal{E})$ is also a solution of Ernst's equation (this is called a Ehlers's transformation for any real number c).

8.3.8 The Kerr–Schild Metric and Two-Black-Hole States

The Kerr solution has been presented in standard Boyer–Lindquist coordinates $(t_{BL}, \phi_{BL}, r, \theta)$. Being a generalization of the Schwarzschild coordinates to the rotating case, these coordinates are singular on the event horizon. For numerical purposes, it is quite often suitable not to use Boyer–Lindquist coordinates, but to go to the Kerr–Schild form, which is a generalization of the Eddington–Finkelstein coordinates. It is possible to cover all, or part, of the interior of a single black hole with a time-independent slicing. However, doing so seems

to require that we give up the maximal-slicing condition. To cover the interior of the black hole, we need a slicing that passes smoothly through the event horizon. A convenient way to generate such solutions is to begin with the metric in standard ingoing-null coordinates. If we want to consider a rotating black hole, then we use the Kerr geometry in Kerr coordinates defined previously.

Coordinate systems attached to physical observers are generically regular at the horizon, but lack the important practical property of stationarity. If, besides regularity, we impose the additional requirements of a stationary metric and space-like foliation, we obtain what we call horizon-adapted coordinates. A comprehensive family of those systems for the Kerr spacetime can be obtained with the following transformation from the standard Boyer–Lindquist (BL) coordinates to the new coordinates (T, ϕ, r, θ)

$$d\phi = d\phi_{\text{BL}} + (a/\Delta) dr \quad (8.206)$$

$$dT = dt_{\text{BL}} + \left[\frac{1+Y}{1+Y-Z} - \frac{1-Z^k}{1-Z} \right] dr, \quad (8.207)$$

where $Y = a^2 \sin^2 \theta / \varrho^2$, $Z = 2Mr / \varrho^2$, and k is a nonnegative integer that parameterizes the family (natural units are used throughout). All members of the family are regular at the horizon; hence, the algebraically simplest choice ($k = 1$) is preferred. This corresponds to the so-called Kerr–Schild (KS) form of the Kerr metric given by the coordinate transformation

$$d\phi = d\phi_{\text{BL}} + (a/\Delta) dr \quad (8.208)$$

$$dT = dt_{\text{BL}} + \frac{2Mr dr}{\Delta}. \quad (8.209)$$

With this choice, the line element becomes

$$\begin{aligned} ds^2 = & -(1-Z) dT^2 - 2aZ \sin^2 \theta dT d\phi + 2Z dT dr \\ & +(1+Z) dr^2 - 2a(1+Z) \sin^2 \theta dr d\Phi + \varrho^2 d\theta^2 \\ & + \sin^2 \theta [\varrho^2 + a^2(1+Z) \sin^2 \theta] d\phi^2. \end{aligned} \quad (8.210)$$

The regularity at the horizon, which is located at the largest root of the equation ($\Delta = 0$), is manifest.

The regularization of the horizon introduces two new nonzero metric elements, as well as a nonvanishing shift vector component β^r . However, this additional algebraic complexity should not be much of a concern for relativistic integration algorithms that must be designed to handle a general metric. On the other hand, it is a fact that considerable intuition and mathematical tools have been obtained in the simpler frozen star form of the Kerr metric. This background work can still be used by transforming geometric quantities back and forth. For stationary accretion patterns, this process is entirely straightforward.

For $k = 1$, the nonzero components of the lapse, shift, and three-metric are then given by

$$\alpha = \frac{\varrho}{\sqrt{\varrho^2 + 2Mr}} \quad (8.211)$$

$$\beta^r = \frac{2Mr}{\varrho^2 + 2Mr} \quad (8.212)$$

$$\gamma_{rr} = 1 + \frac{2Mr}{\varrho^2} \quad (8.213)$$

$$\gamma_{r\phi} = - \left[1 + \frac{2Mr}{\varrho^2} \right] a \sin^2 \theta \quad (8.214)$$

$$\gamma_{\theta\theta} = \varrho^2 \quad (8.215)$$

$$\gamma_{\phi\phi} = \left[r^2 + a^2 + \frac{2Mr}{\varrho^2} a^2 \sin^2 \theta \right] \sin^2 \theta. \quad (8.216)$$

The inverse components are given by

$$\beta_r = 2Mr/\varrho^2 \quad (8.217)$$

$$\beta_\phi = -2aMr \sin^2 \theta / \varrho^2 \quad (8.218)$$

$$\gamma^{rr} = \frac{\Sigma^2}{\varrho^2(\varrho^2 + 2Mr)} \quad (8.219)$$

$$\gamma^{\theta\theta} = 1/\varrho^2 \quad (8.220)$$

$$\gamma^{r\phi} = a/\varrho^2 \quad (8.221)$$

$$\gamma^{\phi\phi} = 1/\varrho^2 \sin^2 \theta. \quad (8.222)$$

Cartesian coordinate components can be obtained from these via the standard Kerr–Schild coordinate transformations

$$x = (r \cos \phi - a \sin \phi) \sin \theta, \quad y = (r \sin \phi + a \cos \phi) \sin \theta, \quad z = r \cos \theta. \quad (8.223)$$

This yields the implicit definition of r from

$$r^4 - r^2(x^2 + y^2 + z^2 - a^2) - a^2z^2 = 0, \quad (8.224)$$

with $r > 0$ and $r = 0$ on the disk described by $z = 0$ and $x^2 + y^2 < a^2$.

The outward unit normal \mathbf{s} on the two-surface $r_H = \text{const.}$ is obtained from $s_i = (\alpha, 0, 0)$ with $\gamma^{ik} s_i s_k = 1$

$$s_i = \left(\varrho \sqrt{\frac{\varrho^2 + 2Mr}{\Sigma^2}}, 0, 0 \right) \quad (8.225)$$

$$s^i = \left(\frac{1}{\varrho} \sqrt{\frac{\Sigma^2}{\varrho^2 + 2Mr}}, \frac{a}{\varrho} \sqrt{\frac{\varrho^2 + 2Mr}{\Sigma^2}}, 0 \right). \quad (8.226)$$

The unit time-like normal to Σ_t is then deduced from the values of the lapse function and the shift vector

$$n^\mu = \left(\sqrt{\varrho^2 + 2Mr}/\varrho, 0, -2Mr/\varrho\sqrt{\varrho^2 + 2Mr}, 0 \right) \quad (8.227)$$

and

$$n_\mu = \left(-\varrho/\sqrt{\varrho^2 + 2Mr}, 0, 0, 0 \right). \quad (8.228)$$

Finally, the extrinsic curvature can be calculated

$$K_{rr} = 2M(a^2 \cos^2 \theta - r^2)(\varrho^2 + Mr)/\varrho^5 \sqrt{\varrho^2 + 2Mr} \quad (8.229)$$

$$K_{r\theta} = 2a^2 Mr \sin \theta \cos \theta / \varrho^3 \sqrt{\varrho^2 + 2Mr} \quad (8.230)$$

$$K_{r\phi} = aM(r^2 - a^2 \cos^2 \theta) \sin^2 \theta \sqrt{\varrho^2 + 2Mr} / \varrho^5 \quad (8.231)$$

$$K_{\theta\theta} = 2Mr^2 / \varrho \sqrt{\varrho^2 + 2Mr} \quad (8.232)$$

$$K_{\theta\phi} = -2a^3 M \sin^3 \theta \cos \theta / \varrho^3 \sqrt{\varrho^2 + 2Mr} \quad (8.233)$$

$$K_{\phi\phi} = r + a^2 M(a^2 \cos^2 \theta - r^2) \sin^2 \theta / \varrho^4. \quad (8.234)$$

In Kerr–Schild coordinates (t, x, y, z) the metric is given by

$$g_{\mu\nu} = \eta_{\mu\nu} + 2Hl_\mu l_\nu, \quad (8.235)$$

where $\eta_{\mu\nu}$ is the Minkowski metric, and l_μ is a null-vector with respect to both the full metric and the Minkowski metric, $g^{\mu\nu}l_\mu l_\nu = 0 = \eta^{\mu\nu}l_\mu l_\nu$. From this we obtain the lapse function, the shift vector and the spatial metric

$$\alpha = 1/\sqrt{1 + 2Hl^t l^t} \quad (8.236)$$

$$\beta^i = -\frac{2Hl^t l^i}{1 + 2Hl^t l^t} \quad (8.237)$$

$$\gamma_{ik} = \delta_{ik} + 2Hl_i l_k. \quad (8.238)$$

For a black hole of mass M and angular momentum $M\mathbf{a}$ at rest at the origin, H and l_μ are given by

$$H = \frac{Mr^3}{r^4 + (\mathbf{a} \cdot \mathbf{x})^2} \quad (8.239)$$

$$l_\mu = (1, \mathbf{l}) \quad (8.240)$$

$$\mathbf{l} = \frac{r\mathbf{x} - \mathbf{a} \times \mathbf{x} + (\mathbf{a} \cdot \mathbf{x})\mathbf{a}/r}{r^2 + \mathbf{a}^2}, \quad (8.241)$$

where

$$r^2 = \frac{\mathbf{x}^2 - \mathbf{a}^2}{2} + \sqrt{\frac{(\mathbf{x}^2 - \mathbf{a}^2)^2}{4} + (\mathbf{a} \cdot \mathbf{x})^2}. \quad (8.242)$$

For a nonrotating black hole, $\mathbf{a} = 0$, H has a pole at the origin, whereas for a rotating black hole H has a ring singularity.

Harmonic Coordinates

Harmonic time-slicing is integral to some hyperbolic formulations of general relativity, and a time-independent harmonic slicing of the Kerr geometry does exist. The harmonic time-slicing condition is $\square t = 0$, which can be written

$$\frac{1}{\sqrt{-g}} \partial_\mu [\sqrt{-g} g^{t\mu}] = 0. \quad (8.243)$$

This equation is satisfied by using the coordinate choice

$$t = V - r + 2M \ln \left| \frac{2M}{r - r_-} \right|, \quad \phi = \Phi. \quad (8.244)$$

For $k = 1$, the nonzero components of the lapse, shift, and three-metric are then given by

$$\alpha^{-2} = 1 + \frac{2Mr}{\varrho^2} \left(\frac{r + r_+}{r - r_-} \right) + \frac{r_+^2 + a^2}{\varrho^2} \left(\frac{2M}{r - r_-} \right) \quad (8.245)$$

$$\beta^r = \alpha^2 \frac{r_+^2 + a^2}{\varrho^2} \quad (8.246)$$

$$\beta^\phi = -\alpha^2 \frac{a}{\varrho^2} \frac{2M}{r - r_-} \quad (8.247)$$

$$\gamma_{rr} = \left[2 - \left(1 - \frac{2Mr}{\varrho^2} \right) \frac{r + r_+}{r - r_-} \right] \frac{r + r_+}{r - r_-} \quad (8.248)$$

$$\gamma_{r\phi} = -\left[1 + \frac{2Mr}{\varrho^2} \frac{r + r_+}{r - r_-} \right] a \sin^2 \theta \quad (8.249)$$

$$\gamma_{\theta\theta} = \varrho^2 \quad (8.250)$$

$$\gamma_{\phi\phi} = \left[r^2 + a^2 + \frac{2Mr}{\varrho^2} a^2 \sin^2 \theta \right] \sin^2 \theta. \quad (8.251)$$

For the harmonic slicing, the hypersurface is space-like only outside the Cauchy horizon at $r > r_-$.

8.4 Rotational Energy and the Four Laws of Black Hole Evolution

Roger Penrose and Stephen Hawking showed 30 years ago that, according to general relativity, any object that collapses to form a black hole will go on to collapse to a singularity inside the black hole. This means that there are strong gravitational effects on arbitrarily short distance scales inside a black hole. On short distance scales, we certainly need to use a quantum theory to describe the collapsing matter. The presence of a singularity in the classical theory also means that once we go

Table 8.2. The four laws of black hole evolution are in strict one-to-one correspondence with the four laws of thermodynamics

Law	Thermodynamic systems	Black Holes
Zero	$T = \text{const}$ in thermal equilibrium	surface gravity κ_H is const on Horizon of stationary BHs
First	$dE = T dS + dW$	$dM_H = (\kappa_H/8\pi) dA_H + \Omega_H dJ_H$
Second	entropy always increases, $\delta S \geq 0$ for all processes	area of BH always increases, $\delta A_H \geq 0$ for all processes involving BHs
Third	$T = 0$ can be never reached	$\kappa_H = 0$ can be never reached

sufficiently far into the black hole, we can no longer predict what will happen. It is hoped that this failure of the classical theory can be cured by quantizing gravity as well.

Using quantum field theory, Hawking has shown that a black hole will radiate thermally. That is, if we study quantum matter fields on a classical black hole background, we find that, when the matter fields are initially in the vacuum (that is, there is no matter falling into the black hole), there is a steady stream of outgoing radiation, which has a temperature determined by its mass and charge.

This is an extremely startling discovery; classically, no radiation can escape from a black hole, but if we quantize the matter fields, we find there is steady flux of radiation coming out of the black hole! This outgoing radiation decreases the mass of the black holes, so eventually the black hole will disappear. The temperature goes up as the black hole gets smaller (unlike most things, which cool off as they lose energy), so the black hole will disappear abruptly, in a final flash of radiation.

There is an *analogy between the classical laws governing black holes, and the laws of thermodynamics*. But thermodynamics is just an approximate description of the behavior of large groups of particles, which works because the particles obey statistical mechanics (a branch of quantum theory). Since black holes have a nonzero temperature, the classical laws of black holes are the laws of thermodynamics applied to black holes. Hence, there must be some more fundamental description of the classical laws governing black holes in terms of statistical mechanics.

8.4.1 Surface Gravity and Angular Velocity of the Horizon

Near the horizon, we can Taylor expand the redshift factor

$$\alpha(r_+) = 0, \quad r_+ = M + \sqrt{M^2 - a^2} \quad (8.252)$$

$$\alpha_H^2 = \frac{\varrho^2}{\Sigma^2} \Delta \simeq \left(1 - \frac{a^2 \sin^2 \theta}{2Mr_+}\right) \left(\frac{r_+ - M}{Mr_+}\right) (r - r_+). \quad (8.253)$$

This surface is still spherical, but the rotation slightly deforms it

$$ds_H^2 = \varrho_H^2 d\theta^2 + R_H^2 d\phi^2 = \gamma_{AB} dx^A dx^B \quad (8.254)$$

$$\varrho_H^2 = r_+^2 + a^2 \cos^2 \theta = 2M_H r_H - a^2 \sin^2 \theta \quad (8.255)$$

$$R_H = \frac{\Sigma_H}{\varrho_H} \sin \theta = \frac{2Mr_+ \sin \theta}{\sqrt{r_+^2 + a^2 \cos^2 \theta}}. \quad (8.256)$$

The circumference at the equator is given by

$$C_H = 2\pi R_H = 4\pi M \quad (8.257)$$

independent of the angular momentum J_H ! Consider now the two-space spanning the null surface. The surface of the horizon can be computed

$$\begin{aligned} A_H &= \int_0^\pi \int_0^{2\pi} \sqrt{g_{\theta\theta} g_{\phi\phi}} d\phi d\theta \\ &= 2\pi \int_0^\pi \sqrt{\sin^2 \theta \Sigma^2} d\theta \\ &= 4\pi (r_+^2 + a^2) = 8\pi M r_+. \end{aligned} \quad (8.258)$$

Surface Gravity

Let u^μ be the velocity field of a stationary observer. Then its four-acceleration is given by

$$a_g^\mu = (u^\varrho \nabla_\varrho) u^\mu. \quad (8.259)$$

For the static observer this yields (see the TOV equation)

$$a_{g\mu} = \nabla_\mu \alpha, \quad (8.260)$$

where $\alpha^2 = \varrho^2 \Delta / \Sigma^2$ is the redshift factor in the Kerr geometry. With this quantity we form a scalar acceleration

$$a_g = \sqrt{-a_{g\mu} a_g^\mu} = \frac{1}{\alpha} \sqrt{-(\nabla_\mu \alpha)(\nabla^\mu \alpha)}. \quad (8.261)$$

This quantity diverges for $r \rightarrow r_+$, as shown, for example, by its expression for the Schwarzschild geometry

$$a_{g\mu} = \frac{GM}{r^2(1-2GM/r)} \nabla_\mu r, \quad (8.262)$$

i.e.

$$a_g = \frac{GM}{r^2 \sqrt{1-2GM/r}}. \quad (8.263)$$

We therefore renormalize the acceleration by means of the redshift factor

$$\kappa = \alpha a_g . \quad (8.264)$$

This quantity, which is called **surface gravity of the horizon**, now remains finite at the horizon and its value is given for the Kerr geometry

$$\kappa_H = \frac{r_+ - M}{2Mr_+} .$$

(8.265)

A fiducial observer (**ZAMO**) would measure a diverging force in the vicinity of the hole, $a_g = -\nabla \ln \alpha \mapsto \infty$. By renormalization of this force (at time basis t) one gets the **surface gravity** of a black hole

$$\mathbf{g}_H \mapsto \alpha_H \mathbf{a}_{gH} = -\kappa_H \mathbf{n}, \quad \kappa_H = c^2 \frac{r_+ - M}{2Mr_+} .$$

(8.266)

\mathbf{n} is the normal vector to the horizon, $\mathbf{n} = \nabla \alpha / |\nabla \alpha| = \mathbf{e}_r$.

Zeroth Law: Both, surface gravity, κ_H , and rotation of the hole, Ω_H , are constant at the horizon (i.e. independent of the poloidal angle θ). The surface of the black hole acts like a rigid rotator.

8.4.2 First Law of Black Hole Dynamics

The total amount of mass-energy of a black hole, $M_H c^2$, consists as in the case of a rotating star of two parts. Therefore, we consider the surface $A_H = 8\pi M_H r_+$, and apply

$$A_H = 4\pi(r_+^2 + a^2) = 4\pi r_+^2 + 4\pi \frac{J_H^2}{M_H^2} = \frac{A_H^2}{16\pi M_H^2} + 4\pi \frac{J_H^2}{M_H^2} . \quad (8.267)$$

Solving this relation for M_H , we get the mass as a function of surface and angular momentum

$$M_H = M_H(A_H, J_H) = \sqrt{\frac{A_H}{16\pi} + 4\pi \frac{J_H^2}{A_H}} .$$

(8.268)

In this way, we found a separation of the total energy into two contributions, whereas the second term corresponds to the rotational energy. One can show that the surface gravity, κ_H , and the angular velocity of the hole, Ω_H , act as intensive variables, while A_H and J_H are the corresponding extensive variables

$$\frac{\kappa_H}{8\pi} = \left(\frac{\partial M_H}{\partial A_H} \right)_{J_H} , \quad \Omega_H = \left(\frac{\partial M_H}{\partial J_H} \right)_{A_H} . \quad (8.269)$$

So, the **First law** of black hole dynamics holds

$$dM = \frac{\kappa_H}{8\pi} dA_H + \Omega_H dJ_H \quad (8.270)$$

for the transition of a black hole from one stationary state to another. This is in full analogy to the first law of thermodynamics if the surface of the hole can be interpreted as an *entropy* and the surface gravity, κ_H , stands for a **temperature**. The corresponding proportionality constants cannot be derived classically.

Hawking found in 1974 that a black hole indeed emits thermal radiation (called **Hawking radiation**) as if the horizon would have a temperature, T_H , given by

$$T_H = \frac{\hbar}{2\pi k_B c} \kappa_H \simeq 6.17 \times 10^{-8} K \frac{M_\odot}{M_H}. \quad (8.271)$$

This is known as the **Hawking temperature**. Therefore, one can assign an entropy to black holes that is proportional to its surface

$$S_H = \frac{k_B}{4\hbar} A_H = \frac{k_B}{4} \frac{\text{Horizon Surface}}{(\text{Plancklength, } 1.6 \times 10^{-33} \text{ cm})^2} \quad (8.272)$$

with ($\Lambda_P = \sqrt{\hbar G/c^3}$ as the Planck length). In case of a nonrotating black hole, this leads to extremely large values for the entropy

$$S_H = \frac{16\pi k_B}{4\hbar} M_H^2 = 1.05 \times 10^{77} k_B (M_H/M_\odot)^2. \quad (8.273)$$

This entropy is 10^{19} times $S_\odot (\simeq 10^{58} k_B)$ for stellar black holes. Supermassive black holes have such small temperatures that the entropy cannot sufficiently decrease via emission of Hawking radiation. In principle, it is still possible to extract angular momentum, i.e. rotational energy. Then, one would arrive at a hole with the same entropy (same surface area) as the incident hole (reversible process). This corresponds to a mass of

$$M_{\text{irr}} = \sqrt{\frac{A_H}{16\pi}} = \frac{1}{2} \sqrt{r_+^2 + a^2}, \quad (8.274)$$

called the **irreducible mass** of a black hole. Irreducible mass, surface area and entropy, S_H , are equivalent expressions for black holes. The first law can be formulated as

$$dM_H = T_H dS_H + \Omega_H dJ_H, \quad (8.275)$$

with temperature, T_H , and black hole angular velocity, Ω_H , as intensive variables

$$T_H = \left(\frac{\partial M_H}{\partial S_H} \right)_{J_H}, \quad \Omega_H = \left(\frac{\partial M_H}{\partial J_H} \right)_{S_H}. \quad (8.276)$$

Hence, the total energy also satisfies

$$M = M(S_H, J_H) = \sqrt{\frac{\hbar}{4\pi k_B} S_H + \frac{\pi k_B}{\hbar} \frac{J_H^2}{S_H}}. \quad (8.277)$$

The Holographic Principle

Under normal circumstances, a vacuum is a space in which there is no matter. But at the quantum level, the vacuum is full of particles and antiparticles that constantly appear and disappear. The Heisenberg uncertainty principle allows these virtual particles and antiparticles to emerge from the vacuum for a brief moment and disappear back into the vacuum again without violating the energy conservation law. According to Hawking, if a particle/antiparticle pair is created near the event horizon of a black hole, gravity will pull one of the particles into the hole permanently, while the other particle (or antiparticle) can escape, or be radiated, from the black hole.

The typical Hawking radiation temperature from solar-mass-sized black holes is close to absolute zero, and radiation becomes fainter as the temperature decreases. Though of fundamental importance in physics, Hawking radiation is very hard to observe directly from space. One curious feature about Hawking radiation is that the temperature is inversely proportional to the mass of the black hole. Thus, the only black holes that might render detectable radiation would be primordial **mini-holes** that may have formed shortly after the Big Bang. Such black holes would have a mass of 10^{15} grams, but a size smaller than an atom. The possibility of detecting such mini-holes, however, is uncertain.

In 1976, Bill Unruh of the University of British Columbia showed that an accelerated observer would experience a similar heat bath of photons around him or her, due also to the existence of an event horizon. The temperature of the heat bath follows the same Hawking temperature formula, except that instead of the gravitational force, it is proportional to the magnitude of the observer's acceleration. Although the Unruh effect induced by acceleration is not precisely the Hawking effect from black holes, it nevertheless shares many common characteristics. It is therefore an intriguing idea that the Hawking effect could be studied using violent acceleration in the laboratory setting, since the temperature associated with the Unruh effect can be much higher if the observer is intensely accelerated.

Chen at SLAC theorized that it should be possible to detect the Unruh radiation emitted by electrons that are accelerated by ultra-intense lasers. One major challenge with detecting Unruh radiation is that enormous accelerations are required to produce sufficient radiation. For example, one would need to accelerate a particle over 10^{20} meters per second squared to generate a temperature of 1 K. It turns out that state-of-the-art lasers can deliver pulses of less than a picosecond (one-trillionth of a second) with petawatts (10^{15} watts) of power. These technologies can in principle accelerate electrons over 10^{25} times the acceleration due to the gravity on Earth's surface, or 10^{28} m/sec², more than two orders of magnitude higher than previous experimental proposals.

The **Holographic Principle**, yet unproven, states that there is a maximum amount of information content held by regions adjacent to any surface. Therefore, counter-intuitively, the information content inside a room depends not on the volume of the room but on the area of the bounding walls. The principle derives from the idea that the Planck length, the length-scale where quantum mechanics begins to dominate classical gravity, is one side of an area that can hold only about one bit of information.

The limit was first postulated by physicist Gerard 't Hooft in 1993 [391]. It can arise from generalizations that the information held by a black hole is determined not by its enclosed volume but by the surface area of its event horizon. The term holographic arises from a hologram analogy where three-dimensional images are created by projecting light through a flat screen.

8.4.3 Rotational Energy of Astrophysical Black Holes

Consider that the total energy $M_H c^2$ is not homogeneous in first order in extensive variables A_H and J_H , moreover this is true in order 1/2. Therefore, the black hole mass holds due to the Euler theorem for homogeneous functions (*Smarr formula*)

$$M_H = \frac{\kappa_H}{4\pi} A_H + 2\Omega_H J_H = 2T_H S_H + 2\Omega_H J_H. \quad (8.278)$$

For a black hole embedded in a material environment, there is an additional contribution by the nonvanishing energy-stress tensor, $T_{\mu\nu}$. Total mass, M^∞ , and total angular momentum J^∞ as measured in the asymptotical flat region satisfy

$$M^\infty = \frac{\kappa_H}{4\pi} A_H + 2\Omega_H J_H - \int_{\Sigma} (2T_v^\mu - T\delta_v^\mu) k^\nu d\sigma_\mu, \quad (8.279)$$

$$J^\infty = J_H + \int_{\Sigma} T_v^\mu m^\nu d\sigma_\mu. \quad (8.280)$$

The integration extends over a space-like hypersurface Σ that intersects the horizon in a 2D surface.

After this considerations, the *rotational energy* of a rotating black hole holds

$$E_{\text{rot}}/c^2 = M_H - M_{\text{irr}} = \sqrt{M_{\text{irr}}^2 + \frac{J_H^2}{4M_{\text{irr}}^2}} - M_{\text{irr}} \quad (8.281)$$

or

$$E_{\text{rot}} = M_H c^2 \left[1 - \sqrt{\frac{1}{2}(1 + \sqrt{1 - a_*^2})} \right] \simeq 5 \times 10^{54} a_*^2 M_{H,8} \text{ watts}, \quad (8.282)$$

where $M_{H,8}$ is the mass of the black hole in units of 100 million solar masses. The maximum value of this rotational energy is then

$$E_{\text{rot,max}} = (1 - \sqrt{1/2}) M_H c^2 = 0.29 M_H c^2, \quad (8.283)$$

i.e. 29% of the total energy of an extremely fast rotating black hole sticks in rotation! Only for slowly rotating objects, the summation over energies is linear, with rotational energy

$$E_{\text{rot}} \simeq \frac{1}{2} I_H \Omega_H^2, \quad I_H = M_H r_+^2. \quad (8.284)$$

In fast rotating black holes, there is a greater amount of energy as compared to supermassive rotators with the same mass. Hence, a black hole disposes of a second energy channel besides accretion energy. If this energy is extracted within the lifetime of an AGN that harbors a central supermassive black hole, then this luminosity is comparable to the accretion power. This process is analogous to the decay of rotational energy in neutron star physics in the pulsar state. This analogy hints for two classes of AGN: One class gets the luminosity from accretion, the other class from rotation.

8.4.4 On the Second and Third Laws of Black Hole Dynamics

The analogy between surface A_H and entropy S_H suggests that the surface is not allowed to decrease via any process. Of course, this is the meaning of the *irreducible* mass. This is a consequence of a consideration of time dependent horizons [107].

The FIDO velocity field $U_{\text{FIDO}} = (\xi + \omega m)/\alpha$ becomes a null field at the horizon

$$l = \xi + \Omega_H m, \quad l^\mu = \frac{dx^\mu}{dt}, \quad l^\mu l_\mu = 0, \quad (8.285)$$

with a renormalization of universal time t . Then, l is a generator field of the 3D horizon and has a geodetic property (outgoing photons on the horizon). Even a non-stationary horizon is generated by such a null geodesics field l . The parameter t has a similar meaning as in the Kerr geometry. The space intersection of the horizons represents a kind of compact bubble that moves with t and changes in form. In general, this bubble A_H does not satisfy axisymmetry but there is

$$l = \partial_t + v^A \frac{\partial}{\partial x^A}, \quad A = 2, 3, \quad \nabla_l l = \kappa_H l. \quad (8.286)$$

Generally speaking, the bubble has a time-dependent metric

$$ds_{|\Sigma_H}^2 = \gamma_{AB}(t) dx^A dx^B. \quad (8.287)$$

$\gamma = \det \gamma_{AB}$ defines a surface element $d\Sigma_H = \sqrt{\gamma} dx^2 dx^3$, which has an expansion

$$\Theta_H = \frac{1}{2} \gamma^{AB} \frac{d\gamma_{AB}}{dt} \quad (8.288)$$

and shear

$$2\sigma_{AB}^H = \frac{d\gamma_{AB}}{dt} - \Theta_H \gamma_{AB}, \quad \sigma_A^{H,A} = 0. \quad (8.289)$$

Therefore, one arrives at the time evolution of the surface

$$\frac{D d\Sigma_H}{dt} = \Theta_H d\Sigma_H. \quad (8.290)$$

Expansion and shear fulfill geometrical identities, the so-called *Raychoudhuri equations*

$$\frac{d\Theta_H}{dt} - \kappa_H \Theta_H = -8\pi D \quad (8.291)$$

$$8\pi D = \frac{1}{2} \Theta_H^2 + \sigma_{AB}^H \sigma^{H,AB} + R_{\alpha\beta} l^\alpha l^\beta \quad (8.292)$$

$$\frac{d\sigma_{AB}^H}{dt} - (\kappa_H - \Theta_H) \sigma_{AB}^H = +(2\sigma_{AC}^H + \gamma_{AC} \Theta_H) \sigma_B^{H,C} - \mathcal{E}_{AB}^H, \quad (8.293)$$

with \mathcal{E}_{AB}^H as tidal forces at the horizon.

Especially, the growth of the surface elements can be computed to

$$\ln \left(\frac{d\Sigma_1}{d\Sigma_0} \right) = \int_{t_0}^{t_1} \Theta_H dt. \quad (8.294)$$

For a transition of two stationary state, one gets with (8.291)

$$\delta(d\Sigma_H) = \frac{8\pi}{\kappa_H} d\Sigma_H \int_{t_0}^{t_1} \left(\frac{1}{16\pi} \Theta_H^2 + \frac{1}{8\pi} \sigma_{AB}^H \sigma^{H,AB} + \frac{1}{8\pi} R_{\alpha\beta} l^\alpha l^\beta \right) dt. \quad (8.295)$$

Together with Einstein's equation, the change of the horizon's surface is given by

$$\delta A_H = \oint \frac{8\pi}{\kappa_H} d\Sigma_H \int_{t_0}^{t_1} \left(\frac{1}{16\pi} \Theta_H^2 + \frac{1}{8\pi} \sigma_{AB}^H \sigma^{H,AB} + T_{\alpha\beta} l^\alpha l^\beta \right) dt \geq 0. \quad (8.296)$$

The first term describes dissipation from expansion and shear of horizon generators. The second term depicts the energy flow into the hole. Hence, the second law of black hole dynamics satisfies:

In any classical accretion process and in any interaction between matter or radiation with the black hole, the horizon area, A_H , can never decrease with time.

These relations state explicitly that a stationary horizon is not allowed to reveal expansion and shear

$$\Theta_H = 0 = \sigma_{AB}^H \sigma^{H,AB}, \quad R_{\alpha\beta} l^\alpha l^\beta = 0. \quad (8.297)$$

Besides this, the second law implies that in a possible collision of two black holes – which may happen in the center of a galaxy – the surface area of the resulting merged black hole *always exceeds* the sum of the two separate progenitor black holes. So, the following relation can be deduced

$$\begin{aligned} M_H & \left(M_H + \sqrt{M_H^2 - a_H^2} \right) \\ & \geq M_1 \left(M_1 + \sqrt{M_1^2 - a_1^2} \right) + M_2 \left(M_2 + \sqrt{M_2^2 - a_2^2} \right). \end{aligned} \quad (8.298)$$

The efficiency of the energy gain is then in turn

$$\epsilon = \frac{M_1 + M_2 - M_H}{M_1 + M_2} < \frac{1}{2}. \quad (8.299)$$

Two colliding nonrotating black holes with equal mass result in $\epsilon \leq 1 - 1/\sqrt{2} = 0.293$.

The explicit form of the surface gravity κ_H (8.266) (temperature) suggests a **third law**:

The surface gravity, κ_H , can never reach zero via a finite number of operations at a stationary state; the faster the black hole rotates, the harder it is to reach the extreme Kerr solution.

8.5 Time Evolution of Black Holes

Astrophysical black holes grow by merging processes and by accretion of vast amount of gas, for example, in the nuclei of galaxies. As a consequence, the mass and angular momentum of black holes will change over the lifetime of the Universe. In this section, we discuss two aspects of this problem: the growth of black holes by accretion of mass and angular momentum, and the formation of new black holes by means of merger processes. This latter event will produce gravitational waves which will be detected in the near future by a new generation of laser interferometer gravitational wave detectors. The coalescence and merger of binary black holes is expected to be one of the primary gravitational radiation to be detected by interferometric gravitational wave detectors. For large binary separations, post-Newtonian approximations can be used to model the binary inspiral with good accuracy [78]. For small binary separations, when finite size and nonlinear effects become more important, only numerical relativity simulations are expected to provide exact wave forms.

8.5.1 Quasistationary Evolution of Accreting Black Holes

The characteristic parameters of rotating black holes are not only mechanical variables, but at the same time thermodynamical ones. For the energy, entropy and angular momentum are the independent quantities, it is however possible to choose temperature and entropy, or angular velocity and angular momentum as independent variables [313]. In this way, one can study evolutionary tracks in the state space, as, for example, adiabatic processes, defined by

$$dS_H = 0, \quad dM_H = \Omega_H dJ_H, \quad (8.300)$$

or processes characterized by an equilibrium between entropy production and extraction of rotational energy

$$-dM_H = T_H dS_H = -\frac{1}{2} \Omega_H dJ_H. \quad (8.301)$$

Black holes mainly evolve through accretion and merging processes. By accretion a black hole gains angular momentum and is speeding up [55]. This process is very similar to the spin-up of accreting neutron stars in LMXBs, which finally produces the millisecond pulsar. Under the assumption that gas accretes from the marginally stable orbit, this gas will bring in a specific angular momentum and a specific internal energy, so that mass and angular momentum of the black hole will grow by the amount

$$dM_H = E_{\text{in}} dM_0, \quad dJ_H = L_{\text{in}} dM_0, \quad r_{\text{in}} = r_{\text{ms}}. \quad (8.302)$$

This provides us an evolution for the specific angular momentum of the black hole with increasing mass

$$\frac{d(a/M)}{d \ln M_H} = \frac{1}{M_H} \frac{L_{\text{in}}}{E_{\text{in}}} - 2a_*, \quad \frac{dM_H}{dM_0} = E_{\text{in}}, \quad (8.303)$$

where $E_{\text{in}} = E_{\text{in}}(M, a_*)$, and $L_{\text{in}} = L_{\text{in}}(M, a_*)$. By integrating this equation, we obtain the evolution of the specific angular momentum and mass, with initial condition $a(M_i) = 0$,

$$a_* = \sqrt{\frac{2}{3}} \frac{M_i}{M_H} \left(4 - \sqrt{18M_i^2/M_H^2 - 2} \right), \quad (8.304)$$

$$M_0 - M_{0i} = 3M_i \left(\sin^{-1}(M_H/3M_i) - \sin^{-1}(1/3) \right), \quad (8.305)$$

where $1 \leq M_H/M_i \leq \sqrt{6}$. In particular, in the limit $a_* \mapsto 1$ we find

$$M_0 - M_{0i} = 3M_i \left(\sin^{-1}(\sqrt{2/3}) - \sin^{-1}(1/3) \right) + \sqrt{3} \left(M_H - \sqrt{6}M_i \right), \quad (8.306)$$

for $M_H/M_i \geq \sqrt{6}$. M_H is the total mass of the hole, M_0 denotes the accreted mass, and M_i, M_{0i} the initial value. According to this, a nonrotating black hole with initial mass M_i and initial specific angular momentum $a_* = 0$ will grow towards an extreme Kerr hole, $a = M$, by accreting a mass $\Delta M_0 = 1.8464 M_i$, while its total mass changes by $\Delta M = (\sqrt{6} - 1)M_i = 1.4459M_i$. For $M_H = \sqrt{6}M_i$, we find $S_H = 3S_i$ with initial entropy S_i , and therefore with an entropy change $\Delta S = 2S_i$. This leads to an expansion of the horizon radius by the amount $\sqrt{6}/2 = 1.2247$. For the irreducible mass $M_{\text{irr}} = M_H/\sqrt{2} = \sqrt{3}M_i$, one finds therefore $\Delta M_{\text{irr}} = (\sqrt{3} - 1)M_i = 0.7321M_i$, corresponding to a change in the rotational energy $\Delta E_{\text{rot}} = (\sqrt{6} - \sqrt{3})M_i = 0.7174M_i$. The rotational energy amounts to 29% of the total mass M_H , however, 49% comes from the accreted mass ΔM .

This would obviously be in contrast to the third law of black hole mechanics, since extreme Kerr is reached by accretion of a finite amount of mass. A more careful analysis shows, however, that there is no contradiction. We can write the above evolution equations as

$$\frac{d \ln M_H}{dt} = \frac{\dot{M}_{\text{acc}}}{M_H} \frac{E_{\text{in}}}{c^2} \quad (8.307)$$

$$\frac{da_*}{dt} = \frac{\dot{M}_{\text{acc}}}{M_H} \left[j_0(a) - 2a_* e_0(a) \right], \quad (8.308)$$

where $j_0 = c j_{\text{in}} / GM_H$, $e_0 = E_{\text{in}} / c^2$, and $a_* = a/M$ is the dimensionless Kerr parameter. This shows that both mass evolution and angular momentum evolution are given by the accretion time-scale, called the *Salpeter time*,

$$\boxed{t_S = \frac{M_H}{\dot{M}_{\text{acc}}} = \frac{1}{\dot{m}} \frac{M_H}{\dot{M}_{\text{Edd}}} = 4 \times 10^7 \text{ yrs} \frac{\epsilon_H}{0.1\dot{m}}.} \quad (8.309)$$

Accretion from the marginally stable orbit satisfies the third law of black hole mechanics, since $j_0(a) - 2ae_0(a) \rightarrow 0$ for $a \rightarrow 1$ (i.e. in the extreme Kerr case). Including Poynting flux losses from the ergosphere, given by \mathcal{P} , the above equations are then changed to

$$\frac{d \ln M_H}{dt} = \frac{\dot{M}_{\text{acc}}}{M_H} \frac{E_{\text{in}}}{c^2} - \frac{\mathcal{P}}{M_H c^2} \quad (8.310)$$

$$\frac{d J_H}{dt} = \dot{M}_{\text{acc}} L_{\text{in}} - \frac{\mathcal{P}}{\Omega_F}. \quad (8.311)$$

Ω_F denotes the angular frequency of magnetic field lines penetrating the ergosphere. Poynting flux emerging from the ergosphere will extract angular momentum and therefore slow down the rotation of the black hole. In the Blandford–Znajek model (see Sect. 8.6), the Poynting flux follows from the expression

$$\mathcal{P} \simeq \frac{1}{8} \frac{B_H^2 r_H^4}{c} \Omega_F (\Omega_H - \Omega_F), \quad (8.312)$$

where B_H is the magnetic field strength at the horizon. With this, the evolution of the specific angular momentum follows from

$$\begin{aligned} \frac{da}{dt} &= \frac{\dot{M}_{\text{acc}}}{M_H} [j_0(a) - 2a e_0(a)] \\ &+ 2a \frac{\mathcal{P}}{M_H c^2} - \frac{\Omega_H}{8} (1 - \Omega_F / \Omega_H) \frac{r_H c^2}{GM_H} \frac{B_H^2 r_H^3}{M_H c^2}. \end{aligned} \quad (8.313)$$

The angular momentum loss is then effectively driven by the rotation time-scale, which is, however, prolonged by a huge factor given by the the black hole energy contained in the horizon volume compared to the total magnetic energy in this volume. Under realistic conditions, magnetic energy is only a tiny fraction of the total black hole energy. For this reason, accretion will always win, and the black holes end up near the extreme Kerr state.

The above arguments, therefore, show that accreting black holes are always driven towards rapidly rotating black holes. Black holes in interaction with magnetic fields could indeed find an equilibrium state with a specific angular momentum near the maximal one. For this reason, black holes found in quasars and other active galactic nuclei are expected to be rapidly rotating objects, similar to the neutron stars found in low-mass X-ray binary systems.

8.5.2 Merging of Black Holes

Binary black hole systems are formed in centers of galaxies, when the cores of two galaxies merge. Such processes are expected to occur frequently in the early Universe, when galaxies form by merging from smaller entities. Black holes will spiral to the common center of the newly formed galaxy by means of dynamical friction. When they come into the range of a few hundred Schwarzschild radii, their further evolution is triggered by the emission of gravitational waves. As a result, the orbits of the two supermassive black holes will tighten by spiraling inwards. This binary black hole problem is still largely unsolved.

The general interest is how the two black holes orbit each other, spiral inwards and eventually merge to form a new black hole. In this process gravitational waves are emitted which should be detectable with future wave detectors (LIGO, VIRGO, GEO and LISA). In the early phase of inspiral, post-Newtonian methods can be applied (Blanchet et al. [77]). But the final evolution has to be treated within full general relativity. This three-dimensional two-black-hole problem has not yet been generally solved. There are two nasty inherent problems which inhibit a simple numerical treatment:

- gauge freedom of Einstein's equations: gauge freedom refers to the freedom to choose coordinates. This is in general not practicable for the entire spacetime in advance. In the coordinate construction process, one has to avoid the formation of coordinate singularities.
- the appearance of a singularity in the case of black hole formation.

Significant progress has been made in the last few years in the dynamical evolution of binary black hole scenarios even in 3D. All these attempts are still facing stability problems that are not completely understood. One reason for the origin of these problems is the handling of the black hole singularities. Simulations of black holes are severely complicated by the presence of singularities. Computers are not built to handle singularities. One method is based on excising the black hole interior. Black hole excision is currently considered the most promising approach to avoid singularities in dynamical simulations. For a discussion of these aspects see, e.g. Baumgarte and Shapiro [62].

Fundamental Equations of Black Hole Mergers

As in the stationary case, spacetime is decomposed into (3+1) dimensions, a procedure which is generally called (3+1)-split of the line element

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ik}(dx^i + \beta^i dt)(dx^k + \beta^k dt). \quad (8.314)$$

This is the generalization of the axisymmetric decomposition. This *foliation* of four-dimensional spacetime into hypersurfaces Σ is characterized in terms of the lapse function α and a shift vector β . γ is now the intrinsic metric of the hypersurface Σ . The dynamical fields are the three-metric γ_{ik} and its extrinsic curvature K_{ik} , both

depending on space and time t . The time evolution of these fundamental quantities is then given in terms of the Arnowitt–Deser–Misner (ADM) formalism [429] (see Sect. 2.8)

$$\partial_t \gamma_{ik} = -2\alpha K_{ik} + D_i \beta_k + D_k \beta_i \quad (8.315)$$

$$\begin{aligned} \partial_t K_{ik} &= -D_i D_k \alpha + \alpha(R_{ik} - 2K_{im} K_k^m + K K_{ik}) \\ &\quad + \beta^m D_m K_{ik} + K_{im} D_k \beta^m + K_{mk} D_i \beta^m \end{aligned} \quad (8.316)$$

The extrinsic curvature K_{ik} acts as a kind of momenta for the metric elements. The lapse function α and the shift vector β have to satisfy the *constraint equations* on each hypersurface

$$R - K_{ik} K^{ik} + K^2 = 0 \quad (8.317)$$

$$D^m (K_{im} - g_{im} K) = 0. \quad (8.318)$$

R_{ik} is the Ricci tensor of the hypersurface, R its trace and K the trace of the extrinsic curvature. D_i denotes the covariant derivative derived from the three-metric γ . Usually, one starts with conformally flat spaces [87]

$$\gamma_{ik} = \psi^4 \delta_{ik}, \quad K_{ik} = \psi^{-2} \bar{K}_{ik}. \quad (8.319)$$

Head-on Collisions

The simplest problem which can be treated is a head-on collision of two black holes, the so-called grazing collision of two black holes (Fig. 8.13, upper panel). In these simulations, two black holes of different masses start from well within the ISCO and are boosted toward each other with a certain impact parameter. The two black holes start out with separate marginally trapped surfaces forming the apparent horizon. In the course of the evolution, a single marginally trapped surface appears which surround the inner trapped surfaces (Fig. 8.13, lower panel). The apparent horizon is defined by a kind of minimal surface equation, and does not evolve continuously, rather a new minimal surface appears at a new location. The shading of the surfaces is a measure for the Gaussian curvature of the surfaces. Alcubierre et al. [28] were able to achieve evolution to about $30 M$. This is important since the lowest quasinormal mode for a Schwarzschild black hole has a period of about $17 M$, which sets the approximate scale for the expected gravitational wavelengths in the ring-down phase.

8.6 Geodesics in the Kerr Geometry

Near black holes in the center of galaxies, stars move on geodesics of the Kerr metric. Similarly, photons escaping from accretion disks near black holes propagate along null geodesics. Such photons get redshifted by the strong gravitational field of the black hole, undergo Doppler shifts, and their trajectories are distorted by lensing effects. For the understanding of spectra emitted by accretion disks around neutron stars and black holes, one needs a relativistic treatment of the photon transport on curved spacetimes.

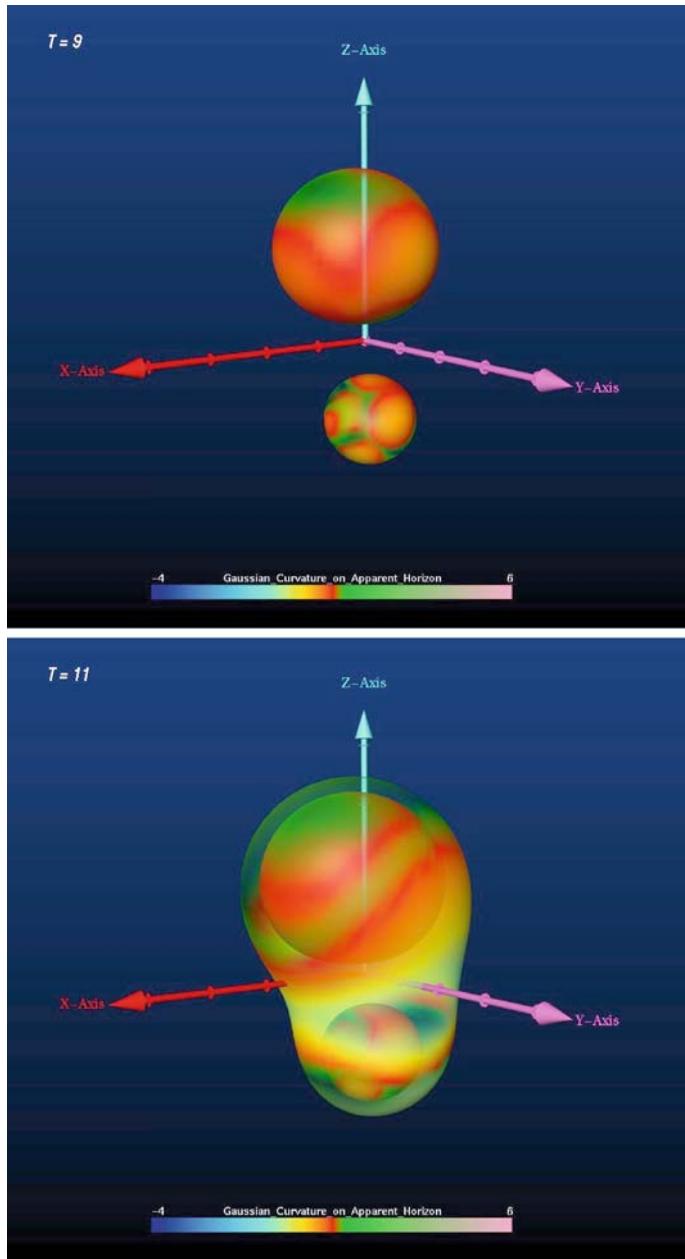


Fig. 8.13. Two black holes in head-on collision (top): the merger of apparent horizons (bottom) in the simulations of Alcubierre et al. [28]. Finally, a new horizon has been formed which surrounds the old ones. This configuration will relax towards a Kerr black hole by means of emission of gravitational waves. The final black hole will only have two hairs: the mass and the spin. Image Credit: AEI Golm

8.6.1 Direct Integration of Geodesics Equations

The geodesics equation

$$\frac{dx^\mu}{d\lambda} = p^\mu \quad (8.320)$$

$$\frac{dp^\mu}{d\lambda} = -\Gamma_{\varrho\sigma}^\mu p^\varrho p^\sigma \quad (8.321)$$

can be directly integrated with the affine parameter λ . These data are assembled into an eight-dimensional vector $X = [x^\mu, p^\mu]$ whose evolution with the parameter λ will be integrated with some leapfrog or Runge–Kutta scheme. As an example of this direct integration, we show the result of shooting photons towards a Schwarzschild black hole in Fig. 8.14 and towards a Kerr hole along the equatorial plane in Fig. 8.15.

In contrast to the nonrotating case, the geodesics are also dragged along by the rotation of the black hole (Fig. 8.15). Near the horizon, photon trajectories are dragged into corotation with the horizon.

When photons start under some inclination with respect to the rotational axis, trajectories are wound up near the horizon (Fig. 8.16).

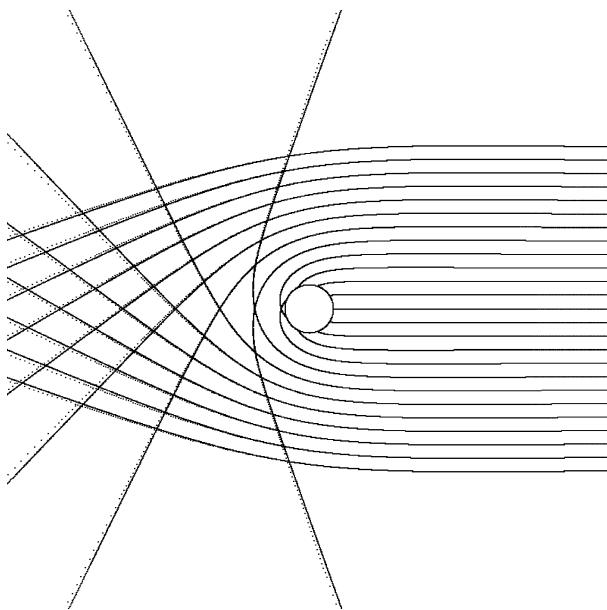


Fig. 8.14. Direct integration of null geodesics starting at some distance along the equatorial plane for a nonrotating black hole, $a = 0$. Figure provided by B. Zink (LSW Heidelberg)

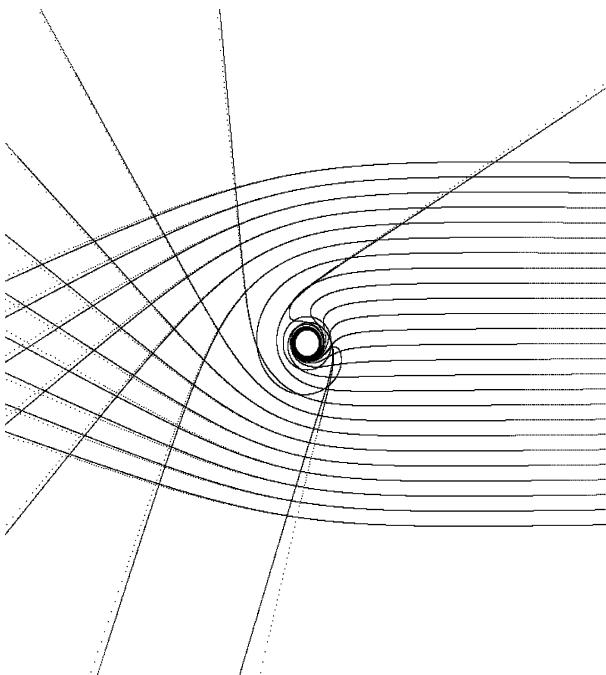


Fig. 8.15. Direct integration of null geodesics starting at some distance away along the equatorial plane for a rotating Kerr black hole. Frame-dragging is so strong near the horizon that the photon trajectories are dragged into the rotational direction. Far away from the horizon, frame-dragging has no influence on the trajectories. Figure provided by B. Zink (LSW Heidelberg)

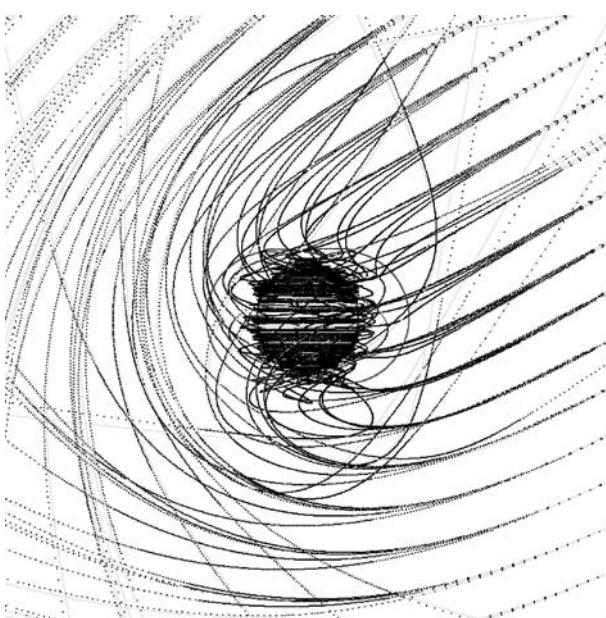


Fig. 8.16. Direct integration of null geodesics starting at some distance at a given inclination with respect to the rotation axis for a rotating Kerr black hole. Figure provided by B. Zink (LSW Heidelberg)

8.6.2 Geodesics in the Equatorial Plane

For the discussion of the motion in the equatorial plane we can rely on similar methods as in the Schwarzschild case (Sect. 8.2.2).

Integral of Motions

The equations of motion for test particles (stars and photons) in the Kerr geometry are completely integrable. Generally, energy E , angular momentum L and the rest mass are constants of motion in any stationary and axisymmetric space-time

$$E = -p_\alpha k^\alpha = -p_t = -g_{tt} p^t - g_{t\phi} p^\phi \quad (8.322)$$

$$L = p_\alpha m^\alpha = p_\phi = g_{t\phi} p^t + g_{\phi\phi} p^\phi \quad (8.323)$$

$$m^2 = p^2. \quad (8.324)$$

This is a consequence of the Killing equation $\nabla_{(\alpha} \xi_{\beta)} = 0$, which means then

$$-(p^\alpha \nabla_\alpha) E = \xi_\beta (p^\alpha \nabla_\alpha) p^\beta + p^\alpha p^\beta \nabla_{(\alpha} \xi_{\beta)} = 0. \quad (8.325)$$

The Radial Equation

As in the case of the discussion of the geodesics in the Schwarzschild geometry, we start from the energy integral

$$2\mathcal{L} = -g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}, \quad (8.326)$$

restricted to motions in the equatorial plane

$$2\mathcal{L} = - \left(1 - \frac{2M}{r}\right) \dot{t}^2 - \frac{4aM}{r} i \dot{\phi} + \frac{r^2}{\Delta} \dot{r}^2 + \left[r^2 + a^2 + \frac{2a^2 M}{r}\right] \dot{\phi}^2. \quad (8.327)$$

Overdots denote differentiation with respect to an affine parameter. From this expression we obtain for the momenta

$$-p_t = \left(1 - \frac{2M}{r}\right) i + \frac{2aM}{r} \dot{\phi} = E \quad (8.328)$$

$$p_r = \frac{r^2}{\Delta} \dot{r} \quad (8.329)$$

$$p_\phi = -\frac{2aM}{r} i + \left[r^2 + a^2 + \frac{2a^2 M}{r}\right] \dot{\phi} = L. \quad (8.330)$$

The corresponding Hamiltonian is given by

$$\begin{aligned}\mathcal{H} &= p_t \dot{t} + p_r \dot{r} + p_\phi \dot{\phi} - \mathcal{L} \\ &= -\frac{1}{2} \left(1 - \frac{2M}{r}\right) \dot{t}^2 - \frac{2aM}{r} \dot{t} \dot{\phi} + \frac{r^2}{2\Delta} \dot{r}^2 \\ &\quad + \frac{1}{2} \left[r^2 + a^2 + \frac{2a^2M}{r}\right] \dot{\phi}^2.\end{aligned}\quad (8.331)$$

Since \mathcal{H} is independent of t , we find

$$\begin{aligned}2\mathcal{H} &= -\left[\left(1 - \frac{2M}{r}\right) \dot{t} + \frac{2aM}{r} \dot{\phi}\right] \dot{t} + \frac{r^2}{\Delta} \dot{r}^2 \\ &\quad + \left[\left(r^2 + a^2 + \frac{2a^2M}{r}\right) \dot{\phi} - \frac{2aM}{r} \dot{t}\right] \dot{\phi}.\end{aligned}\quad (8.332)$$

This can be written as

$$2\mathcal{H} = -E \dot{t} + L \dot{\phi} + \frac{r^2}{\Delta} \dot{r}^2 = -\delta_1 \quad (8.333)$$

where δ_1 is an integration constant, which can be chosen as $\delta_1 = -1$ for time-like geodesics and $\delta_1 = 0$ for null geodesics. With these expressions we can solve for the velocity components in terms of the conserved quantities E and L

$$\dot{t} = \frac{1}{\Delta} \left[\left(r^2 + a^2 + \frac{2a^2M}{r}\right) E - \frac{2aM}{r} L \right] \quad (8.334)$$

$$\dot{\phi} = \frac{1}{\Delta} \left[\left(1 - \frac{2M}{r}\right) L + \frac{2aM}{r} E \right]. \quad (8.335)$$

When we insert this into the equation (8.333), we obtain the **radial equation**

$$\boxed{r^3 \dot{r}^2 = r^3 E^2 + 2M(aE - L)^2 + r(a^2 E^2 - L^2) + \delta_1 r \Delta.} \quad (8.336)$$

Time-like Geodesics and Circular Orbits

For time-like geodesics we may write this in order of decreasing powers in r as

$$r^3 \dot{r}^2 = r^3 E^2 - r \Delta - r(L^2 - a^2 E^2) + 2M(aE - L)^2. \quad (8.337)$$

E is now the total energy per unit mass of a particle (or star) and L the angular momentum per unit mass. This radial equation can directly be obtained from the conservation laws and the normalization of the four momentum, $p^2 = -m^2$.

The radial equation reduces to the Schwarzschild form in the case $a = 0$ (see Sect. 8.2.2)

$$\left(\frac{dr}{d\tau}\right)^2 = E^2 - V_S^2 \quad (8.338)$$

$$\frac{d\phi}{d\tau} = \frac{L}{r^2}, \quad (8.339)$$

with the effective potential defined as

$$V_S^2 = \left(1 - \frac{2M}{r}\right) \left(1 + \frac{L^2}{r^2}\right). \quad (8.340)$$

The possible orbits in Schwarzschild have been discussed in Sect. 8.2.3.

In the case of Kerr, we cannot transform to a simple effective potential. For this reason, we introduce, as in the Newtonian case, the variable $u = 1/r$ and write the radial equation as

$$u^{-4}\dot{u}^2 = 2M(L - aE)^2 u^3 - (L^2 - a^2 E^2) u^2 - (a^2 u^2 - 2Mu + 1) + E^2. \quad (8.341)$$

We are now interested in circular orbits, $\dot{u} = 0$, for given values of E and L . For circular orbits, the above cubic polynomial will have a double root. This is easily calculated to be the case for

$$2Ml^2 u^3 - (l^2 + 2alE) u^2 - (a^2 u^2 - 2Mu + 1) + E^2 = 0 \quad (8.342)$$

and

$$3Ml^2 u^2 - (l^2 + 2alE) u - (a^2 u - M) = 0, \quad (8.343)$$

where we have introduced the reduced angular momentum

$$l = L - aE. \quad (8.344)$$

These two equations can be combined to give

$$E^2 = 1 - Mu + Ml^2 u^3 \quad (8.345)$$

and

$$2alEu = l^2(3Mu - 1)u - (a^2 u - M). \quad (8.346)$$

We can eliminate the energy E from these two equations and combine them into a quartic equation for l^2

$$\begin{aligned} & u^2[(3Mu - 1)^2 - 4a^2Mu^3]l^4 \\ & - 2u[(3Mu - 1)(a^2 u - M) - 2a^2u(Mu - 1)]l^2 \\ & + (a^2 u - M)^2 = 0. \end{aligned} \quad (8.347)$$

The discriminant of this equation is

$$D = 4a^2Mu^3 D_u^2, \quad D_u = a^2u^2 - 2Mu + 1. \quad (8.348)$$

With the introduction of

$$R_{\pm} = 1 - 3Mu \pm 2a\sqrt{Mu^3} \quad (8.349)$$

and the identity

$$(3Mu - 1)^2 - 4a^2Mu^3 = R_+R_- \quad (8.350)$$

we find

$$l^2u^2 = \frac{R_{\pm}D_u - R_+R_-}{R_+R_-} = \frac{1}{R_{\mp}}(D_u - R_{\mp}). \quad (8.351)$$

The last factor can be written in the form

$$D_u - R_{\mp} = u \left(a\sqrt{u} \pm \sqrt{M} \right)^2. \quad (8.352)$$

The two solutions for l can then be written

$$l_{\pm} = -\frac{a\sqrt{u} \pm \sqrt{M}}{\sqrt{uR_{\mp}}}. \quad (8.353)$$

The upper sign applies to retrograde orbits, while the lower sign applies to prograde orbits.

Inserting this solution into the energy equation (8.345), we find

$$E = \frac{1 - 2Mu \mp a\sqrt{Mu^3}}{\sqrt{R_{\mp}}}, \quad (8.354)$$

and the value of L to be associated with this value for E is

$$L = l + aE = \mp \frac{\sqrt{M} \left[a^2u^2 + 1 \pm 2a\sqrt{Mu^3} \right]}{\sqrt{uR_{\mp}}}. \quad (8.355)$$

For the following it is useful to transform back to radii. The energy per unit mass of circular orbits is given as

$$E = \frac{r^2 - 2Mr \mp a\sqrt{Mr}}{r\sqrt{r^2 - 3Mr \mp 2a\sqrt{Mr}}}, \quad (8.356)$$

and the specific angular momentum as

$$L = \mp \frac{\sqrt{Mr}(r^2 - 2a\sqrt{Mr} + a^2)}{r\sqrt{r^2 - 3Mr \mp 2a\sqrt{Mr}}}. \quad (8.357)$$

As stated above, the upper sign applies to retrograde orbits, the lower sign to prograde orbits. We also went back to the geometrical notation for mass and Kerr parameter.

We also can find expressions for the angular velocity Ω of circular orbits

$$\Omega = \frac{d\phi}{dt} = \frac{L - 2Mul}{(r^2 + a^2)E - 2aMlu} \quad (8.358)$$

and for the velocity $v^{(\phi)}$ of the orbit relative to Bardeen observers

$$v^{(\phi)} = \frac{\varpi(\Omega - \omega)}{\alpha} = \left(r^2 + a^2 + \frac{2Ma^2}{r} \right) \frac{\Omega}{\sqrt{\Delta}} - \frac{2aM}{r\sqrt{\Delta}}. \quad (8.359)$$

The relativistic Keplerian angular velocity of circular orbits is therefore given by

$$\Omega_K = \mp \frac{\sqrt{M}}{\sqrt{r^3} \mp a\sqrt{M}}. \quad (8.360)$$

These orbits reach the velocity of light, $E \rightarrow \infty$ for $r = r_{\text{ph}}$, called photon orbit

$$r_{\text{ph}} = 2M \left[1 + \cos \left(\frac{2}{3} \cos^{-1}(-a/M) \right) \right]. \quad (8.361)$$

As we will see in the discussion of null geodesics, such orbits are unstable.

Innermost Stable Orbit (ISCO)

A circular orbit with reciprocal radius u_c is given by some specific values of E and L . With these values chosen, the cubic polynomial on the right-hand side of equation (8.336) will have a double root for $u = u_c$, so we can reduce the radial equation to the form

$$u^{-4}\dot{u}^2 = 2M(L - aE)^2(u - u_c)^2 \left[u + 2u_c - \frac{L^2 - a^2E^2 + a^2}{2M(L - aE)^2} \right]. \quad (8.362)$$

From the expressions for E and L , we find

$$\frac{L^2 - a^2E^2 + a^2}{2Ml^2} = \left[1 + 3a^2u_c^2 \pm 4a\sqrt{Mu_c^3} \right] \frac{1}{2(a\sqrt{u_c} \pm \sqrt{M})^2} \quad (8.363)$$

and

$$u_c - \frac{L^2 - a^2E^2 + a^2}{2Ml^2} = -\frac{D_u}{2(a\sqrt{u_c} \pm \sqrt{M})^2}. \quad (8.364)$$

Accordingly we arrived at the expression for the radial equation

$$\dot{u}^2 = 2Ml^2 u^4 (u - u_c)^2 (u - u_3), \quad (8.365)$$

where

$$u_3 = -u_c + \frac{D_u}{2 \left(a\sqrt{u_c} \pm \sqrt{M} \right)^2} \quad (8.366)$$

defines the reciprocal radius of the orbit of the second kind associated with stable circular orbits of reciprocal radius u_c (see next section).

The condition for the instability of the circular orbit is that u_c is a triple root (see the Schwarzschild case), i.e. $u_* = u_c$. Equation (8.366) therefore leads to the condition

$$4u_c \left(a\sqrt{u_c} \pm \sqrt{M} \right)^2 = D_u = a^2 u_c^2 - 2Mu_c + 1. \quad (8.367)$$

This can be expanded to the form

$$3a^2 u_c^2 + 6Mu_c \pm 8a\sqrt{Mu^3} - 1 = 0, \quad (8.368)$$

or, reverting to the variable r

$$r_{\text{ms}}^2 - 6Mr_{\text{ms}} \pm 8a\sqrt{Mr_{\text{ms}}} - 3a^2 = 0. \quad (8.369)$$

This quartic equation for \sqrt{r} defines the radius of marginal stability (also called the innermost stable circular orbit (ISCO)). It can be solved with standard methods in a parametrized form

$$r_{\text{ms}} = M_H \left(3 + Z_2 \mp \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)} \right)$$

(8.370)

with the parameters

$$Z_1 = 1 + (1 - a_*^2)^{1/3} \left((1 + a_*)^{1/3} + (1 - a_*)^{1/3} \right) \quad (8.371)$$

$$Z_2 = \sqrt{3a_*^2 + Z_1^2}. \quad (8.372)$$

These radii are plotted as a function of the Kerr parameter a in Fig. 8.11.

Besides the limiting case $E^2 \rightarrow \infty$, the **marginally bound orbit** which satisfies $E^2 = 1$ is of a certain interest. This orbit is reached when a particle, at rest at infinity, falls towards the black hole. From the solution for l and the expression for the energy E we find

$$1 = l^2 u^2 = \frac{u}{R_{\mp}} \left(a\sqrt{u} \pm \sqrt{M} \right)^2, \quad (8.373)$$

or

$$R_{\mp} = 1 - 3Mu \mp 2a\sqrt{Mu^3} = u \left[a^2u + M \pm 2a\sqrt{Mu} \right]. \quad (8.374)$$

This equation simplifies to

$$\left[au \pm 2\sqrt{Mu} \right]^2 = 1, \quad (8.375)$$

or to the expression

$$r_{\text{mb}} = 2M \pm a + 2\sqrt{M^2 \pm aM}. \quad (8.376)$$

All these three radii satisfy the inequality $r_{\text{ms}} > r_{\text{mb}} > r_{\text{ph}}$ for all values of the Kerr parameter a . In the limit $a \rightarrow M$, all three radii seem to converge towards the horizon radius $r_+(a = M) = M$. A more careful analysis of this limit by means of

$$a = M(1 - \delta) \quad (8.377)$$

leads to

$$r_+ = M \left[1 + \sqrt{2\delta} \right] + O(\delta^{3/2}). \quad (8.378)$$

Therefore, we obtain the limits given in Table 8.3.

One might expect from the fact that, since all orbital radii tend to the horizon radius for extreme Kerr, the energy of the marginally stable orbit tends to infinity. A direct calculation from the expression for the total energy E gives however

$$E(a = M) = \frac{1 - 2Mu + \sqrt{Mu^3}}{1 - 3Mu + 2\sqrt{Mu^3}} \rightarrow 1/\sqrt{3}, \quad a \rightarrow M. \quad (8.379)$$

This gives the maximum energy per unit mass which a stable circular orbit can have in a Kerr geometry with $a^2 \leq M^2$. The maximum binding energy of this orbit is therefore $E_{\text{bind}} = 1 - 1/\sqrt{3} = 70\%$ of the rest mass. For all Kerr parameters a , the binding energy of the innermost stable orbit is given by the total energy E , $E_{\text{bind}} = 1 - E(a)$. In an accretion process, half of this energy can be liberated in terms of radiation from the disk around the black hole. This high gravitational binding energy is the reason why black holes can so efficiently transform accretion streams into radiation.

Table 8.3. Limiting radii of circular orbits for Schwarzschild and extreme Kerr

Orbit	$a = 0$	$a = M(1 - \delta)$	$a = -M$
r_{photon}	$3M$	$M(1 + \sqrt{8\delta/3})$	$4M$
$r_{\text{mar. bound}}$	$4M$	$M(1 + 2\sqrt{\delta})$	$(3 + 2\sqrt{2})M$
r_{ISCO}	$6M$	$M(1 + [4\delta]^{1/3})$	$9M$

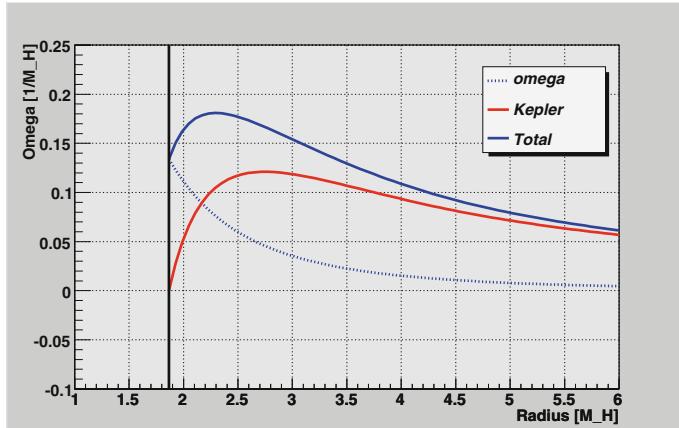


Fig. 8.17. Angular velocity for quasi-Keplerian motion in the equatorial plane for $a_* = 0.5$. The lower solid curve represents the angular velocity determined by the specific angular momentum, the upper curve is the result of both effects, frame-dragging (dotted curve) and angular momentum. The vertical solid line marks the position of the outer horizon

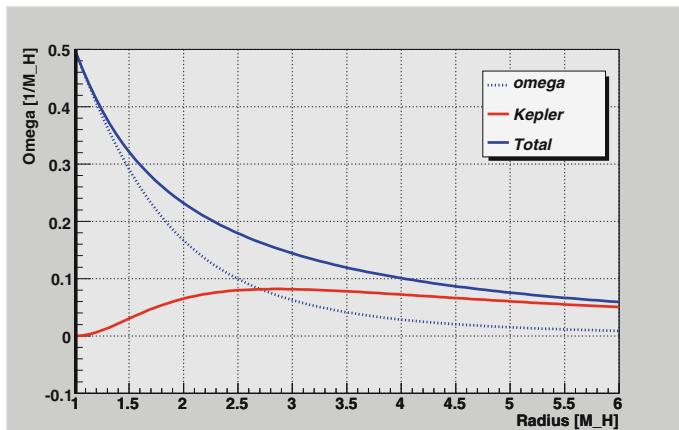


Fig. 8.18. Angular velocity for quasi-Keplerian motion in the equatorial plane for $a_* = 0.9999$. The meaning of the curves is the same as in the previous figure. Now, frame-dragging is the dominant contribution within three gravitational radii

Near the horizon, rotational motion is dominated by frame-dragging. Within the marginal stable orbit, matter falls towards the horizon at constant specific angular momentum λ . Figure 8.17 shows the resulting angular velocity for a low Kerr parameter, $a = 0.5$, given by

$$\Omega(r) = \omega(r) + \frac{\alpha^2(r)}{R^2(r)} \frac{\lambda}{1 - \omega\lambda}. \quad (8.380)$$

Beyond the marginal stable orbit, the specific angular momentum is given by equation (8.355). In the extreme Kerr limit, rotation will be completely dominated by frame-dragging within about four gravitational radii (Fig. 8.18).

This represents an extremely important insight into the properties of rapidly rotating black holes. While in accretion disks, the evolution of the plasma is dictated by the shear of the Keplerian motion, near the horizon this effect is completely negligible and shear will always be dominated by frame-dragging.

8.6.3 Geodesics Including Lateral Motion

So far, we have only discussed motion in the equatorial plane. Including lateral motion of matter would require a fourth constant of motion in order to be able to solve the geodesics equations. It is another miracle of the Kerr geometry that indeed such an integral of motion exists. Carter [105] has demonstrated explicitly the separability of the Hamilton–Jacobi equation and deduced from this the existence of a further integral of motion.

The Fourth Integral of Motion

The above discussed three integrals of motion do not allow us to solve completely for the momenta. In stationary and axisymmetric spacetimes, one can, however, find two additional conserved quantities, based on the existence of a symmetric Killing tensor [411], $K_{\alpha\beta} = K_{\beta\alpha}$, which has the essential property $\nabla(\alpha K_{\beta\gamma}) = 0$, and is defined as

$$K_{\alpha\beta} = 2\varrho^2 l(\alpha n_\beta) - r^2 g_{\alpha\beta}. \quad (8.381)$$

Here, the metric is expressed in terms of the tetrads $\mathbf{l}, \mathbf{n}, \mathbf{m}$

$$-g_{\alpha\beta} = l_\alpha n_\beta + l_\beta n_\alpha - m_\alpha \bar{m}_\beta - m_\beta \bar{m}_\alpha. \quad (8.382)$$

For the Kerr geometry, these tetrads have the following form

$$\mathbf{l} = \frac{1}{\Delta} (r^2 + a^2, \Delta, 0, a) \quad (8.383)$$

$$\mathbf{n} = \frac{1}{2\varrho^2} (r^2 + a^2 2, -\Delta, 0, a) \quad (8.384)$$

$$\mathbf{m} = \frac{1}{\bar{\varrho}\sqrt{2}} (ia \sin \theta, 0, 1, i/\sin \theta) \quad (8.385)$$

with $l^2 = 0 = n^2$ (null vector),

$$\bar{\varrho} = r + ia \sin \theta \quad (8.386)$$

and

$$\mathbf{l} \cdot \mathbf{n} = 1, \quad \mathbf{m} \cdot \bar{\mathbf{m}} = -1. \quad (8.387)$$

We now define two scalars

$$K = K_{\alpha\beta} p^\alpha p^\beta = 2\varrho^2 (\mathbf{p} \cdot \mathbf{l})(\mathbf{p} \cdot \mathbf{n}) - r^2 \mathbf{p}^2 \quad (8.388)$$

$$K = 2\varrho^2 (\mathbf{p} \cdot \mathbf{m})(\mathbf{p} \cdot \bar{\mathbf{m}}) + a^2 \mathbf{p}^2 \cos^2 \theta, \quad (8.389)$$

which are then conserved along geodesic motions due to the Killing equation

$$(p^\alpha \nabla_\alpha) K = 2K_{\beta\gamma} p^\beta (p^\alpha \nabla_\alpha) p^\gamma + p^\alpha p^\beta p^\gamma \nabla_{(\alpha} K_{\beta)\gamma} = 0. \quad (8.390)$$

Therefore, K will have the following form

$$K = \frac{1}{\Delta} (p^t - a\Delta \sin^2 \theta p^\phi)^2 - \frac{\varrho^4}{\Delta} (p^r)^2 - m^2 r^2 \quad (8.391)$$

and

$$K = (a \sin \theta p^t - (r^2 + a^2) \sin \theta p^\phi)^2 + \varrho^4 (p^\theta)^2 + m^2 a^2 \cos^2 \theta. \quad (8.392)$$

On the other hand, we find for E and L_z the two relations

$$a \sin \theta p^t - (r^2 + a^2) \sin \theta p^\phi = aE \sin \theta - L_z / \sin \theta \quad (8.393)$$

and

$$\Delta p^t - a\Delta \sin^2 \theta p^\phi = (r^2 + a^2)E - aL_z. \quad (8.394)$$

With this, we may rewrite the expression for K

$$K = \frac{1}{\Delta} ((r^2 + a^2)E - aL_z)^2 - \frac{\varrho^4}{\Delta} (p^r)^2 - m^2 r^2 \quad (8.395)$$

$$K = (aE \sin \theta - L_z / \sin \theta)^2 + \varrho^4 (p^\theta)^2 + m^2 a^2 \cos^2 \theta. \quad (8.396)$$

When solved for the momenta, this leads to explicit expressions for the poloidal momenta

$$\varrho^4 (p^r)^2 = ((r^2 + a^2)E - aL_z)^2 - \Delta(m^2 r^2 + K) \quad (8.397)$$

$$\varrho^4 (p^\theta)^2 = -(aE \sin \theta - L_z / \sin \theta)^2 + K - m^2 a^2 \cos^2 \theta. \quad (8.398)$$

Separability of the Hamilton–Jacobi Equation

As we have discussed at various places, the geodesic equations follow from a variational principle

$$S[\tau] = \int_0^\tau L ds, \quad \mathcal{L} = -\frac{1}{2} g_{\alpha\beta} \frac{dx^\alpha[\lambda(s)]}{ds} \frac{dx^\beta[\lambda(s)]}{ds}. \quad (8.399)$$

The Euler–Lagrange equations

$$\frac{d}{ds} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} \right) - \frac{\partial \mathcal{L}}{\partial x^\alpha} = 0 \quad (8.400)$$

imply then the geodesic equations (see Sect. 3.5). Via Legendre transformation we obtain the Hamiltonian

$$\mathcal{H}(x, p) = -\dot{x}^\alpha(p) p_\alpha - L(x, \dot{x}(p)) = -\frac{1}{2} g^{\alpha\beta}(x) p_\alpha p_\beta. \quad (8.401)$$

We now consider the above action from the point of view that S characterizes the true orbits. We may compare the value of the action integral for different orbits with the same initial condition, but different end points for the proper time τ . In this sense, the action integral is considered for real trajectories in spacetime as a function of the end point position. This has the consequence that the partial derivatives of the action coincide with the momenta

$$p_\alpha = \frac{\partial S}{\partial x^\alpha}. \quad (8.402)$$

The derivative of the action with respect to proper time is identical to the Hamilton function

$$\frac{\partial S}{\partial \tau} - \mathcal{H} = 0. \quad (8.403)$$

When we replace the momenta we arrive at

$$\boxed{\frac{\partial S}{\partial \tau} - \mathcal{H} \left(x^\alpha; \frac{\partial S}{\partial x^\alpha} \right) = 0.}$$

(8.404)

This represents a partial differential equation of first order which is known under the name *Hamilton–Jacobi equation*.

We consider the action $S[\tau, x^\alpha]$ along geodesics for fixed initial conditions in the Kerr geometry

$$\frac{\partial S}{\partial \tau} = -\frac{1}{2} g^{\alpha\beta} \frac{\partial S}{\partial x^\alpha} \frac{\partial S}{\partial x^\beta}. \quad (8.405)$$

For the Kerr metric, this equation becomes

$$\begin{aligned} 2 \frac{\partial S}{\partial \tau} &= \frac{\Sigma^2}{\varrho^2 \Delta} \left(\frac{\partial S}{\partial t} \right)^2 + \frac{4aMr}{\varrho^2 \Delta} \frac{\partial S}{\partial t} \frac{\partial S}{\partial \phi} \\ &\quad - \frac{\Delta - a^2 \sin^2 \theta}{\varrho^2 \Delta \sin^2 \theta} \left(\frac{\partial S}{\partial \phi} \right)^2 \\ &\quad - \frac{\Delta}{\varrho^2} \left(\frac{\partial S}{\partial r} \right)^2 - \frac{1}{\varrho^2} \left(\frac{\partial S}{\partial \theta} \right)^2. \end{aligned} \quad (8.406)$$

It is convenient to transform this equation into an alternative form

$$2 \frac{\partial S}{\partial \tau} = \frac{1}{\varrho^2 \Delta} \left[(r^2 + a^2) \frac{\partial S}{\partial t} + a \frac{\partial S}{\partial \phi} \right]^2 - \frac{1}{\varrho^2 \sin^2 \theta} \left[a \sin^2 \theta \frac{\partial S}{\partial t} + \frac{\partial S}{\partial \phi} \right]^2 - \frac{\Delta}{\varrho^2} \left(\frac{\partial S}{\partial r} \right)^2 - \frac{1}{\varrho^2} \left(\frac{\partial S}{\partial \theta} \right)^2. \quad (8.407)$$

We now seek for solutions which are separable

$$S = \frac{1}{2} m^2 \tau - Et + L_z \phi + S_r(r) + S_\theta(\theta).$$

(8.408)

The Hamilton–Jacobi equation (8.407) implies therefore

$$\begin{aligned} m^2 \varrho^2 = & \frac{1}{\Delta} [(r^2 + a^2)E - aL_z]^2 - \frac{1}{\sin^2 \theta} (aE \sin^2 \theta - L_z)^2 \\ & - \Delta \left(\frac{\partial S_r}{\partial r} \right)^2 - \left(\frac{\partial S_\theta}{\partial \theta} \right)^2. \end{aligned} \quad (8.409)$$

By means of the identity

$$(aE \sin^2 \theta - L_z)^2 \operatorname{cosec}^2 \theta = (L_z^2 \operatorname{cosec}^2 \theta - a^2 E^2) \cos^2 \theta + (L_z - aE)^2 \quad (8.410)$$

we obtain

$$\begin{aligned} & \left\{ m^2 r^2 - \frac{1}{\Delta} [(r^2 + a^2)E - aL_z]^2 + (L_z - aE)^2 + \Delta \left(\frac{\partial S_r}{\partial r} \right)^2 \right\} \\ & + \left\{ [a^2 m^2 + (L_z^2 \operatorname{cosec}^2 \theta - a^2 E^2)] \cos^2 \theta + \left(\frac{\partial S_\theta}{\partial \theta} \right)^2 \right\} = 0. \end{aligned} \quad (8.411)$$

The expressions in the curly brackets are mere functions of r and θ , respectively. Both brackets have to be constant, and this common constant \mathcal{C} corresponds to the existence of a fourth constant of motion. The separability leads therefore to the two relations

$$\begin{aligned} \Delta \left(\frac{\partial S_r}{\partial r} \right)^2 = & \frac{1}{\Delta} [(r^2 + a^2)E - aL_z]^2 \\ & - [\mathcal{C} + (L_z - aE)^2 + m^2 r^2] \end{aligned} \quad (8.412)$$

$$\left(\frac{\partial S_\theta}{\partial \theta} \right)^2 = \mathcal{C} - (L_z^2 \operatorname{cosec}^2 \theta - a^2 E^2 + m^2 a^2) \cos^2 \theta. \quad (8.413)$$

With the notions of

$$P \equiv (r^2 + a^2)E - L_z \quad (8.414)$$

$$\mathcal{R} \equiv P^2 - \Delta[m^2 r^2 + (L_z - aE)^2 + \mathcal{C}] \quad (8.415)$$

$$\Theta \equiv \mathcal{C} - [a^2(m^2 - E^2) + L_z^2 \operatorname{cosec}^2 \theta] \cos^2 \theta, \quad (8.416)$$

the solutions can be written as

$$S = \frac{1}{2}m^2\tau - Et + L_z\phi + \int^r \frac{\sqrt{\mathcal{R}}}{\Delta} dr + \int^\theta \sqrt{\Theta} d\theta \quad (8.417)$$

with the corresponding momenta

$$p_r = \frac{dS_r}{dr} = \frac{\sqrt{\mathcal{R}}}{\Delta}, \quad p_\theta = \frac{dS_\theta}{d\theta} = \sqrt{\Theta}. \quad (8.418)$$

The equations of motion follow from derivatives with respect to the conserved quantities \mathcal{C} , m^2 , E and L_z

$$\frac{\partial S}{\partial \mathcal{C}} = \frac{1}{2} \int \frac{1}{\Delta\sqrt{\mathcal{R}}} \frac{\partial \mathcal{R}}{\partial \mathcal{C}} dr + \frac{1}{2} \int \frac{1}{\Theta} \frac{\partial \Theta}{\partial \mathcal{C}} d\theta = 0, \quad (8.419)$$

or explicitly

$$\int^r \frac{dr}{\sqrt{\mathcal{R}}} = \int^\theta \frac{d\theta}{\sqrt{\Theta}}. \quad (8.420)$$

Similarly, we find

$$\tau = \int^r \frac{r^2}{\sqrt{\mathcal{R}}} dr + a^2 \int^\theta \frac{\cos^2 \theta}{\sqrt{\Theta}} d\theta \quad (8.421)$$

$$\begin{aligned} t &= \frac{1}{2} \int^r \frac{1}{\Delta\sqrt{\mathcal{R}}} \frac{\partial \mathcal{R}}{\partial E} dr + \frac{1}{2} \int^\theta \frac{1}{\Theta} \frac{\partial \Theta}{\partial E} d\theta \\ &= \int^r \frac{(r^2 + a^2)P}{\Delta\sqrt{\mathcal{R}}} dr + a \int^\theta \frac{L_z - aE \sin^2 \theta}{\sqrt{\Theta}} d\theta \end{aligned} \quad (8.422)$$

$$\begin{aligned} \phi &= -\frac{1}{2} \int^r \frac{1}{\Delta\sqrt{\mathcal{R}}} \frac{\partial \mathcal{R}}{\partial L_z} dr - \frac{1}{2} \int^\theta \frac{1}{\sqrt{\Theta}} \frac{\partial \Theta}{\partial L_z} d\theta \\ &= a \int^r \frac{P}{\Delta\sqrt{\mathcal{R}}} dr + \int^\theta \frac{L_z \operatorname{cosec}^2 \theta - aE}{\sqrt{\Theta}} d\theta. \end{aligned} \quad (8.423)$$

The constant \mathcal{C} is uniquely given by the Killing constant K as follows

$$\mathcal{C} = K - (L_z - aE)^2. \quad (8.424)$$

The above functions \mathcal{R} and Θ , respectively, are identical to the expressions found here. In addition, null geodesics can be treated in the same manner just by setting $m^2 = 0$.

On the basis of these relations, the velocities of the geodesics in the Kerr geometry are explicitly given by the following relations, (the constants of integration are normalized per unit mass, $m^2 = \{1, 0\}$)

$$\varrho^2 \frac{dr}{d\tau} = \sqrt{\mathcal{R}} \quad (8.425)$$

$$= \sqrt{[E(r^2 + a^2) - L_z a]^2 - \Delta [m^2 r^2 + (L_z - aE)^2 + \mathcal{C}]} \quad (8.425)$$

$$\varrho^2 \frac{d\theta}{d\tau} = \sqrt{\Theta} = \sqrt{\mathcal{C} - \cos^2 \theta \left(a^2(m^2 - E^2) + \frac{L_z^2}{\sin^2 \theta} \right)} \quad (8.426)$$

$$\Delta \frac{d\phi}{d\tau} = \frac{L_z}{\sin^2 \theta} \left(1 - \frac{2Mr}{\varrho^2} \right) + \frac{2Mr}{\varrho^2} aE \quad (8.427)$$

$$\alpha^2 \frac{dt}{d\tau} = E - \omega L_z. \quad (8.428)$$

τ is the proper time along the geodesic. Null geodesics just follow from $m \mapsto 0$.

We may summarize the results we obtained for the momenta of geodetic particles

$$p_r = \frac{\varrho^2}{\Delta} p^r = \frac{\sqrt{\mathcal{R}}}{\Delta} \quad (8.429)$$

$$p_\theta = \varrho^2 p^\theta = \sqrt{\Theta} \quad (8.430)$$

$$-p_t = E \quad (8.431)$$

$$p_\phi = L_z. \quad (8.432)$$

Similarly, we obtain for the momenta with respect to Bardeen observers

$$-p_{(t)} = p^{(t)} = \frac{1}{\alpha} (E - \omega L_z) = \alpha p^t \quad (8.433)$$

$$p_{(\phi)} = p^{(\phi)} = \varpi p^\phi = p_\phi / \varpi = L_z / \varpi \quad (8.434)$$

$$p_{(r)} = p^{(r)} = \exp(-\mu_2) p_r \quad (8.435)$$

$$p_{(\theta)} = p^{(\theta)} = \exp(-\mu_3) p_\theta. \quad (8.436)$$

Features of the Lateral Motion

The θ motion can be characterized by means of the equation

$$\varrho^2 \dot{\theta}^2 = f(\mu) = \mathcal{C} + A\mu^2 + B\mu^4 \quad (8.437)$$

for $\mu = \cos \theta$ and

$$A = -(\mathcal{C} + L^2 - a^2[E^2 - 1]), \quad B = -a^2(E^2 - 1). \quad (8.438)$$

This function obviously satisfies

$$\begin{aligned} f(0) &= \mathcal{C}, & f(1) &= -L^2, \\ f'(0) &= 2A, & f'(1) &= 2(2B + A). \end{aligned} \quad (8.439)$$

Motion is only possible for $f(\mu) \geq 0$, a trajectory can reach the axis $\mu^2 = 1$, if and only if $\mathcal{C} = 0$.

If $\mathcal{C} > 0$, we see from $f(0) > 0$ that f has a zero μ_0 in the range $0 < \mu \leq 1$. For $\mathcal{C} > 0$, a particle moves in an oscillatory way, repeatedly crossing the equatorial plane with θ in the range $\theta_0 \leq \theta \leq \pi - \theta_0$, where $\cos \theta_0 = \mu_0$. If $\mathcal{C} = 0$, there is a trivial case with $E^2 - 1 = 0 = L$. Then θ may take a constant value. In general, a motion in the equatorial plane, $\mathcal{C} = 0$, is only stable against small θ motions if $L^2 > a^2(E^2 - 1)$.

In the case $\mathcal{C} < 0$, then $f(0) < 0$ and $f(1) \leq 0$. When $A > 0$, we find a maximum of f at some μ_m with $f(\mu_m) \geq 0$. This maximum can occur if $4B\mathcal{C} \geq A^2$ (which requires $E^2 > 1$), or

$$[\mathcal{C} + L^2 - a^2(E^2 - 1)]^2 + 4a^2(E^2 - 1)\mathcal{C} \leq 0. \quad (8.440)$$

A necessary condition for this inequality to be satisfied is $L^2 + a^2(E^2 - 1) > 0$ for $E^2 - 1 > 0$. The equality occurs if there is a double root at μ_m (corresponding to a conical motion). In general the motion is also oscillatory with $0 < \mu_1 \leq \mu \leq \mu_0 \leq 1$. In this case, the trajectory never crosses the equatorial plane.

These considerations are important for accretion processes. Here one studies only motions in the equatorial plane, at least for cool disks. One has then to show that these motions are stable and that a typical motion always tends to evolve towards the equatorial plane.

Radial Motions

The radial function \mathcal{R} is a polynomial of fourth order in the radius

$$\begin{aligned} \mathcal{R} = & (E^2 - m^2)r^4 + 2Mm^2r^3 + [a^2(E^2 - m^2) - L_z^2 - \mathcal{C}]r^2 \\ & + 2M[\mathcal{C} + (aE - L_z)^2]r - a^2\mathcal{C}. \end{aligned} \quad (8.441)$$

It reduces to the polynomial of third order we have found for radial motions in the equatorial plane, equation (8.336), since the fourth integral of motion vanishes in this case, $p_\theta = 0$.

Positions, where \mathcal{R} vanishes, mark the turning points in radial direction. Since Δ vanishes on the horizon r_+ , we find $\mathcal{R}(r_+) \geq 0$ and $\mathcal{R}(r) \rightarrow (E^2 - 1)r^4$ for $r \rightarrow \infty$. Therefore, we can distinguish between four different cases depending on the signs of $E^2 - 1$ and \mathcal{C} . As an example we may think in terms of the motion of compact stars (e.g. neutron stars) around supermassive black holes.

1. $E > 1$ and $\mathcal{C} > 0$: For sufficiently large energies all the coefficients in the polynomial $\mathcal{R}(r)$ of equation (8.441), except the constant, are nonnegative. From a sketch of this function one can derive the following properties. Since $\mathcal{R}(0) < 0$ and $\mathcal{R}(r_+) \geq 0$, there must be a zero at $r = r_1$, in or on the horizon. Increasing the angular momentum L for fixed energy and Carter constant will change the shape of the curve to a second zero at $r = r_2 \simeq 2M$ and a third one at $r_3 > 2M$. For a star moving inwards from infinity, this means that it will cross the horizon for low angular momentum and will be reflected. However, it cannot recross the horizon – such a star will be trapped within the horizon. For

sufficiently high angular momentum, stars will be reflected outside the horizon to infinity. Any star with $r < r_2$ is trapped and will ultimately cross the horizon. Circular orbits with $r_2 = r_3$ are unstable.

2. $E < 1$ and $\mathcal{C} > 0$: Since $\mathcal{R}(0) < 0$, $\mathcal{R}(M) \geq 0$ and $\mathcal{R}(r) \rightarrow +\infty$ for $r \rightarrow \infty$, there must be at least two real zeros of \mathcal{R} . In the case of two zeros, stars move along trapped orbits. In the case of four zeros, any star with $r_3 < r < r_4$ will be oscillating and move on a bound orbit. The case $r_3 = r_4$ corresponds stable circular orbits discussed previously. Particles in the range $r_1 < r < r_2$ are trapped. A close inspection of the form of $\mathcal{R}(r)$ reveals that spherical orbits with negative angular momentum (retrograde orbits) are at larger radii than for positive angular momenta (prograde orbits). This is the effect of the spin-angular momentum interaction mediated by the gravitomagnetic field.
3. $E > 1$ and $\mathcal{C} < 0$: This case is very similar to case (1). All trapped particles will reach $r = 0$.
4. $E < 1$ and $\mathcal{C} < 0$: In this case we find trapped and bound orbits, but all trapped particles will reach $r = 0$.

In summary, we have found that for $E > 1$ all stars that are initially far from the black hole will be reflected back to infinity. Particles very close to the horizon will be trapped. For $E < 1$, no star can reach infinity – any star is either trapped or bound. For the latter case, stable circular orbits can exist. A negative angular momentum represents a repelling force.

8.6.4 Null Geodesics and Ray-Tracing in Kerr Geometry

Ray-tracing in the Kerr Geometry

The observed spectrum of an accretion disk or the spectrum of emission lines from the inner disk (e.g. Fe lines) is essentially influenced by gravitational redshift, Doppler beaming and lensing effects. While in earlier days, these effects have been taken into account via a transfer function [127], modern approaches use directly ray-tracing in the Kerr geometry based on null geodesics. There are essentially two ways to achieve this goal: (i) by means of direct integration of the geodesics equations; (ii) by implementing fast solvers for the elliptical integrals involved in the solutions of the integrals of motion as discussed above (see, e.g. Fanton et al. [153]). The former method is easily implemented using some leapfrog scheme, the shooting of at least one million photons towards the central object is still very time-consuming and requires implementation on supercomputers (Martocchia, Karas and Matt [272]). The imaging of a radiating torus around a Kerr black hole is shown in Fig. 8.20. The second method is very elegant, however restricted to applications on the Kerr geometry only. It is not known, whether null geodesics are integrable on a general axisymmetric spacetime. We discuss in the following the second method following essentially the pioneering paper by Fanton et al. [95, 153].

The principle of ray-tracing is simple (Fig. 8.19). One shoots about a million photons towards the black hole from an observer's position given by (r_0, θ_0, ϕ_0) ,

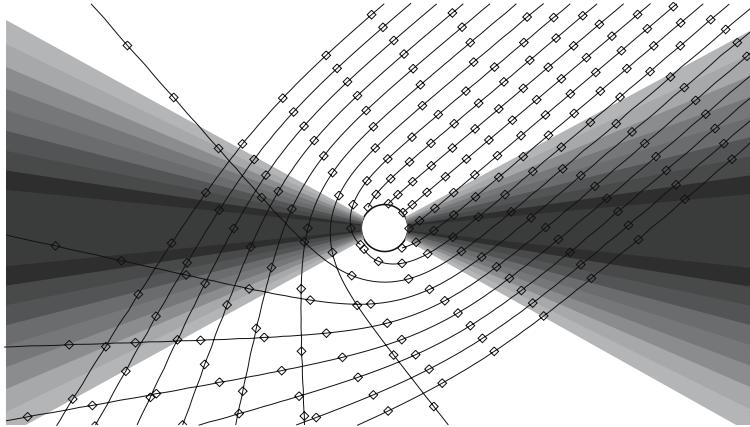


Fig. 8.19. Principle of ray-tracing in the Kerr geometry. The spinning black hole is in the center of the coordinate system, surrounded by an accretion disk with outer radius r_{out} and inner radius ending at the horizon. Photon paths are integrated along geodesic trajectories from a distant observer to the region around the black hole. The positions and momenta are tabulated along the photon path and used to integrate the radiative transfer equation through the accretion disk towards the observer. The photons either terminate at the black hole horizon, or escape to infinity. In the case of optically thick disks, they always terminate at the disk surface located near the equatorial plane. Figure adapted from Schnittman et al. [357]

until they hit an obstacle, e.g. the central accretion disk. The difference between normal ray-tracing and ray-tracing in the Kerr geometry consists in the fact that the light paths are curved geodesics, determined by means of the geodesic relation

$$\int_{r_{\text{em}}}^{r_0} \frac{dr}{\sqrt{\mathcal{R}(r; \lambda, q)}} = \int_{\theta_{\text{em}}}^{\theta_0} \frac{d\theta}{\sqrt{\Theta(\theta; \lambda, q)}} \quad (8.442)$$

with the expressions

$$\mathcal{R}(r; \lambda, q) = r^4 + (a^2 - \lambda^2 - q)r^2 + 2[q + (\lambda - a)^2]r - a^2q \quad (8.443)$$

$$\Theta(\theta; \lambda, q) = q + (a \cos \theta)^2 - (\lambda \cot \theta)^2. \quad (8.444)$$

For null geodesics, the energy E can be absorbed into the scaling so that

$$\lambda = \frac{L_z}{E}, \quad q = \frac{C}{E^2}. \quad (8.445)$$

For ray-tracing, the two parameters λ and q are related to the impact parameters, or to angles α and β on the celestial sphere

$$\alpha = \left(\frac{rp^{(\phi)}}{-p^{(t)}} \right)_{r \rightarrow \infty} = \lambda \operatorname{cosec} \theta_0 \quad (8.446)$$

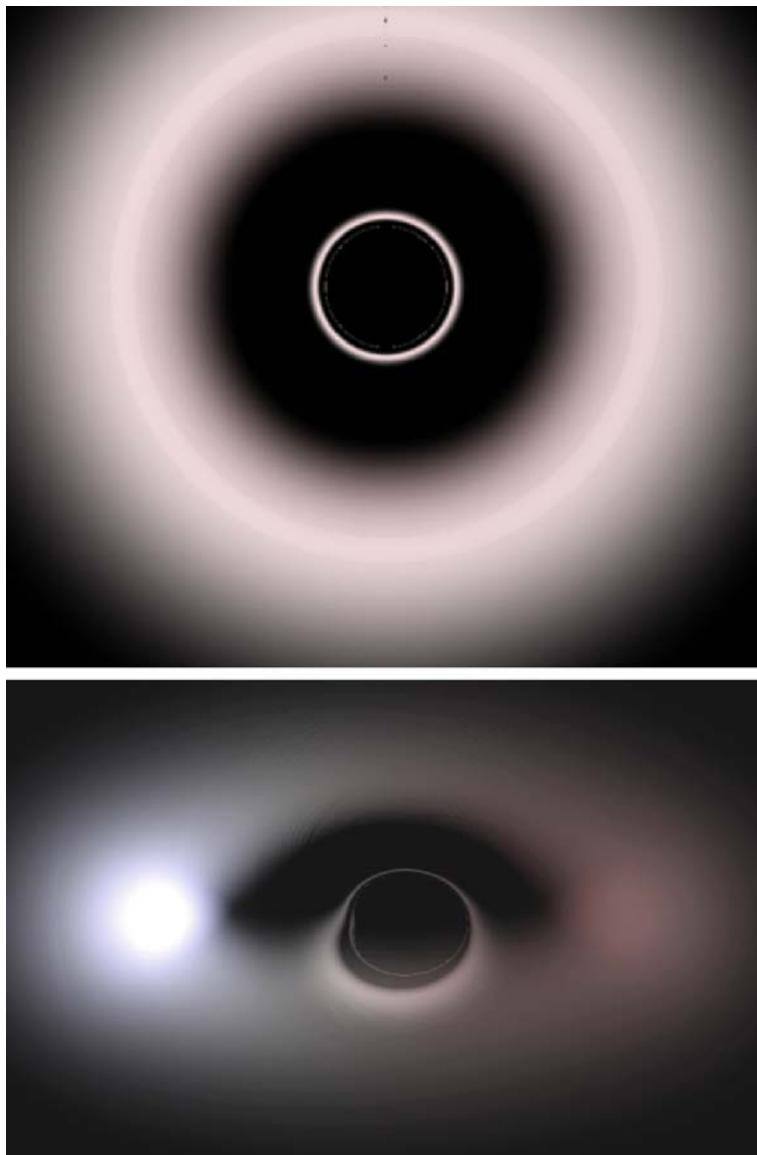


Fig. 8.20. Images formed by Kerr geometry. The *top panel* shows the image of a torus located near the marginal stable orbit of a rapidly rotating black hole. The *lower panel* shows the image of the same torus at an inclination of 60 degrees. At higher inclination, the torus is visible as a bright spot due to Doppler beaming together with some enhanced emission near the horizon (secondary images). Images provided by B. Zink (ZAH, LSW Heidelberg)

and

$$\beta = \left(\frac{rp^{(\theta)}}{-p^{(t)}} \right)_{r \rightarrow \infty} = p_{\theta_0} = \sqrt{q + a^2 \cos^2 \theta_0 - \lambda^2 \cot^2 \theta_0}. \quad (8.447)$$

α is the apparent perpendicular distance of the image from the axis of symmetry, and β is the apparent perpendicular distance of the image from its projection on the equatorial plane. Since photon rays are determined by two integrals of motion, the two impact parameters determine them uniquely.

With $\mu = \cos \theta$ the second equation assumes the form

$$I_\mu = \int \frac{d\mu}{\sqrt{\Theta_\mu}} \quad (8.448)$$

with

$$\Theta_\mu = q - (\lambda^2 + q - a^2)\mu^2 - a^2\mu^4. \quad (8.449)$$

On the equatorial plane we find $\Theta_\mu = q$, while on the rotational axis $\Theta_\mu = -\lambda^2 \leq 0$. We have to distinguish therefore the two cases $q > 0$ and $q < 0$. Only in the former case, $q > 0$, the trajectories can cross the equatorial plane, and they may oscillate around the equatorial plane. This is therefore the case relevant for rays starting from the disk in the equatorial plane. When $q = 0$ the trajectories stay in the equatorial plane and eventually fall into the singularity. In the opposite case, $q < 0$, the trajectories are confined within a cone around the axis, but never cross the equatorial plane. So for the ray-tracing case, we can factorize

$$\Theta_\mu = a^2(\mu^2 + \mu_-^2)(\mu_+^2 - \mu^2) \quad (8.450)$$

with $0 \leq \mu^2 \leq \mu_+^2$ and

$$\mu_\pm^2 = \frac{1}{2a^2} \left[\sqrt{(\lambda^2 + q - a^2)^2 + 4a^2q} \mp (\lambda^2 + q - a^2) \right]. \quad (8.451)$$

In this case, the integral can be written in terms of Jacobian integrals of the first kind

$$\begin{aligned} I_\mu &= -\frac{1}{a\sqrt{\mu_+^2 + \mu_-^2}} F(\arccos(\mu/\mu_+), m_\theta) \\ &= -\frac{1}{a\sqrt{\mu_+^2 + \mu_-^2}} \operatorname{cn}^{-1}(\mu/\mu_+, m_\theta), \end{aligned} \quad (8.452)$$

where

$$m_\theta \equiv \frac{\mu_+^2}{\mu_+^2 + \mu_-^2}. \quad (8.453)$$

Similarly, we consider the integration over r . The equation $R = 0$ has in general four complex solutions. These roots are associated with the turning points of the geodetic motion. Since $R(r = 0) < 0$ and $R(r) > 0$ for $r \rightarrow \infty$, we find at least two real roots, or four of them. One has then the following important cases (Fanton et al. [153]):

1. Four simple roots with $r_1 > r_2 > r_3 > r_4$ with the integral given as

$$\int_{r_1}^r \frac{dr'}{\sqrt{\mathcal{R}}} = \frac{2}{\sqrt{(r_1 - r_3)(r_2 - r_4)}} \operatorname{sn}^{-1} \left(\sqrt{\frac{(r_2 - r_4)(r - r_1)}{(r_1 - r_4)(r - r_2)}}, m_4 \right) \quad (8.454)$$

with

$$m_4 = \frac{(r_2 - r_3)(r_1 - r_4)}{(r_2 - r_4)(r_1 - r_3)}. \quad (8.455)$$

2. Two real solutions r_1 and $r_2 < r_1$ and a double root $r_3 = -(r_1 + r_2)/2$:

$$\begin{aligned} \int \frac{dr}{(r_1 - r_2)\sqrt{(r - r_1)(r - r_2)}} &= -\frac{1}{\sqrt{(r_1 - r_3)(r_2 - r_3)}} \times \\ &\ln \left[\frac{\sqrt{(r - r_1)(r - r_2)}}{r_3 - r} + \frac{r_3^2 + r_1 r_2 + 2r_3 r}{(r_3 - r)\sqrt{(r_1 - r_3)(r_2 - r_3)}} \right] \end{aligned} \quad (8.456)$$

3. Two complex solutions and two real roots (one positive and the other one negative)

$$r_a = u - iw \quad (8.457)$$

$$r_b = u + iw \quad (8.458)$$

$$r_c = -u - v \quad (8.459)$$

$$r_d = -u + v. \quad (8.460)$$

Then the integral can be written as

$$\begin{aligned} \int_r^\infty \frac{dr}{\sqrt{Q_1 Q_2}} &= \frac{1}{w\sqrt{\lambda_1 - \lambda_2}} \\ &\times \left[\operatorname{sn}^{-1}(\sqrt{1 - 1/\lambda_1}, m_2) - \operatorname{sn}^{-1} \left(\sqrt{\frac{1 - 1/\lambda_1}{Q_1(r)}} B_1(r), m_2 \right) \right] \end{aligned} \quad (8.461)$$

with the following abbreviations

$$Q_1(r) = r^2 - 2ur + u^2 + w^2 \quad (8.462)$$

$$Q_2(r) = r^2 + 2ur + u^2 - v^2 \quad (8.463)$$

$$m_2 = \frac{\lambda_1}{\lambda_1 - \lambda_2} \quad (8.464)$$

$$\lambda_{1,2} = \frac{1}{2w^2} \left(\pm \sqrt{(4u^2 + w^2 - v^2)^2 + 4v^2w^2} + (4u^2 + w^2 - v^2) \right) \quad (8.465)$$

$$Q_2 - \lambda_1 Q_1 = (1 - \lambda_1) B_1^2 \quad (8.466)$$

$$B_1(r) = r + u \frac{1 + \lambda_1}{1 - \lambda_1} \quad (8.467)$$

$$B_2(r) = r + u \frac{1 + \lambda_2}{1 - \lambda_2} \quad (8.468)$$

In order to avoid the calculation of the inverse of the Jacobian integrals, the geodetic integral will be written in terms of a trace parameter P

$$\int_r^{r_0} \frac{dr}{\sqrt{\mathcal{R}(r; \lambda, q)}} = P = \int_\theta^{\theta_0} \frac{d\mu}{\sqrt{\Theta_\mu(\mu; \lambda, q)}} \quad (8.469)$$

with

$$\mu = \mu_+ \operatorname{cn} \left(a \sqrt{\mu_+^2 + \mu_-^2} P \pm \Psi_0, m_\theta \right). \quad (8.470)$$

The quantity Ψ_0 is defined as follows

$$\Psi_0 = \operatorname{cn}^{-1} \left(\frac{\mu_{in}}{\mu_+}, m_\theta \right). \quad (8.471)$$

Similarly, the r -integrals are expressed in terms of the trace parameter P

$$r = r(P; r_1, r_2, r_3, r_4), \quad (8.472)$$

so, e.g. in case (i)

$$r = \frac{r_1(r_2 - r_4) - r_2(r_1 - r_4) \operatorname{sn}^2(P \sqrt{(r_2 - r_4)(r_1 - r_3)} / 2 - \xi_0, m_4)}{r_2 - r_4 - (r_1 - r_2) \operatorname{sn}^2(P \sqrt{(r_2 - r_4)(r_1 - r_3)} / 2 - \xi_0, m_4)} \quad (8.473)$$

where

$$\xi_0 = \operatorname{sn}^{-1} \left(\sqrt{\frac{r_2 - r_4}{r_1 - r_4}}, m_4 \right). \quad (8.474)$$

In addition, one can also calculate the azimuthal angle ϕ of the emitting point on the disk which follows from the integral (8.423). Similar expressions follow for the other cases (see Fanton et al. [153]).

A given pixel of the camera determines the two parameters λ and q , and therefore μ_\pm , as well as the four roots of the radial equation. We then calculate the value P_{em} for the disk $\theta = \theta_{\text{em}} \simeq \pi/2$ and get the value r_{em} for each pixel of the camera, depending on the number of roots for the radial equation. Since the emitter is assumed to be axisymmetric, all the physical quantities only depend on θ_{em} and r_{em} .

The Doppler Factor

An important quantity is the redshift factor (or Doppler factor)

$$D = \frac{v_{\text{obs}}}{v_{\text{em}}} = \frac{\hat{p}_{\text{obs}}^t}{\hat{p}_{\text{em}}^t} \quad (8.475)$$

with \hat{p}^a as the photon momenta in the plasma frame. They are related to the photon momenta $p^{(a)}$ in the ZAMO frame over a Lorentz transformation

$$\hat{p}^t = \gamma [p^{(t)} - v^{(i)} p_{(i)}] \quad (8.476)$$

$$\hat{p}^i = p^{(i)} + \frac{1}{v^2} [(\gamma - 1) v^{(j)} p_{(j)} v^{(i)} - \gamma p^{(t)} v^{(i)}]. \quad (8.477)$$

γ is the Lorentz factor of the plasma motion and \mathbf{v} the corresponding three-velocity in the ZAMO frame

$$v^{(r)} = \frac{\exp \mu_r}{\alpha} v^r \quad (8.478)$$

$$v^{(\theta)} = \frac{\exp \mu_\theta}{\alpha} v^\theta \quad (8.479)$$

$$v^{(\phi)} = \frac{R}{\alpha} (\Omega - \omega). \quad (8.480)$$

The ZAMO photon momenta are all proportional to the total energy E , $p^{(a)} = E n^a$,

$$p^{(t)} = \frac{E}{\alpha} (1 - \lambda \omega) \quad (8.481)$$

$$p^{(\phi)} = \frac{E \lambda}{R} \quad (8.482)$$

$$p^{(r)} = \pm \frac{E}{\varrho \Delta} \sqrt{\mathcal{R}_{\lambda,q}/E^2} \quad (8.483)$$

$$p^{(\theta)} = \pm \frac{E}{\varrho} \sqrt{\Theta_{\lambda,q}/E^2} \quad (8.484)$$

with

$$\mathcal{R}_{\lambda,q}/E^2 = r^4 + (a^2 - \lambda^2 - q)r^2 + 2[q + (\lambda - a)^2]r - a^2q \quad (8.485)$$

$$\Theta_{\lambda,q}/E^2 = q + a^2 \cos^2 \theta - \lambda^2 \cot^2 \theta. \quad (8.486)$$

By introducing these quantities into the energy of the photon we get

$$\begin{aligned} \hat{p}^t &= \gamma [p^{(t)} - v^{(r)} p_{(r)} - v^{(\theta)} p_{(\theta)} - v^{(\phi)} p_{(\phi)}] \\ &= \gamma \left[\frac{1 - \lambda \omega}{\alpha} - v^{(r)} \frac{\sqrt{\mathcal{R}}}{\varrho \sqrt{\Delta}} - v^{(\theta)} \frac{\sqrt{\Theta}}{\varrho} - v^{(\phi)} \frac{\lambda}{\varpi} \right]. \end{aligned} \quad (8.487)$$

With this we obtain for the Doppler factor the general expression for any flow field at the position (r, θ)

$$D = \frac{p_{\text{obs}}^{(t)}}{\gamma_E [p^{(t)} - v^{(r)} p_{(r)} - v^{(\theta)} p_{(\theta)} - v^{(\phi)} p_{(\phi)}]_E}. \quad (8.488)$$

The photon propagation only depends on the parameters $\lambda = L_z/E$ and $q = \mathcal{C}/E^2$, the energy of the photons drops out. Since $p^{(t)}$ and $v^{(\phi)}$ are proportional to $1/\alpha$, we multiply the expression with the redshift factor α and cancel the energy

$$D(r, \theta) = \frac{\alpha(r_E, \theta_E)}{\gamma_E [1 - \alpha v^{(r)} n_r - \alpha v^{(\theta)} n_\theta - \lambda \Omega]_E}. \quad (8.489)$$

Obviously, the Doppler factor has to vanish at the horizon, since $\alpha(r_H, \theta) = 0$. The gravitational redshift α will therefore suppress considerably the Doppler factor immediately inside the marginal stable orbit. The Doppler factor will also depend on the poloidal motion of the emitting plasma [303].

In this way, the ray-tracing procedure determines the Doppler factor for each point on the disk surface. The Doppler factor also determines the emission observed by the camera

$$F_v^{\text{Pixel}} = \int_{\text{Pixel}} d\Omega I_v^{\text{obs}} \quad (8.490)$$

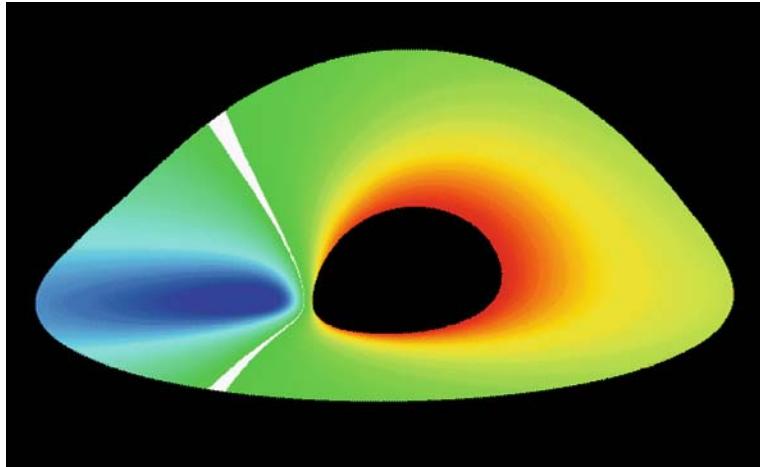


Fig. 8.21. Redshift distribution for geometrically thin disks. Near the horizon gravitational redshift is dominant, $D < 1$. On the left flank, the emission is blueshifted by the Doppler effect, $D > 1$. The white line denotes positions on the disk, where the observed emission is not shifted, $D = 1$

and over the Liouville theorem we obtain

$$F_v^{\text{Pixel}} = \int_{\text{Pixel}} d\Omega D^3 I_v^{\text{em}}. \quad (8.491)$$

In particular for line emission we can make the approximation

$$I_v^{\text{em}} = \epsilon(r) \delta(v^{\text{em}} - v_0), \quad (8.492)$$

where v_0 represents the rest frame frequency of the line. The line emissivity $\epsilon(r)$ has to be prescribed by some suitable physical model. This is usually chosen as some power law, $\epsilon(r) \propto r^{-p}$ for some range $r_{\text{in}} \leq r \leq r_{\text{out}}$, since all physical quantities of a standard disk have some scaling behavior.

The numerical procedure for calculating an image of an axisymmetric disk is therefore straightforward [303]:

- specify the Kerr parameter a , disk parameters $r_{\text{in}}, r_{\text{out}}$, as well as the velocity profiles for the emitting plasma v^r, v^θ and Ω , and the parameters for the location of the screen;
- specify an emissivity law $\epsilon(r)$;
- specify the impact parameters (α, β) ;
- calculate λ and q ;
- evaluate (r_E, θ_E) for given rays;
- evaluate the Doppler factor $D(r_E, \theta_E; \lambda, q)$;
- sum over all pixels to evaluate the flux F_v^{obs} .

In this way one can construct redshift images of the disk (Fig. 8.21), as well as temperature images for standard accretion disks.

Profiles of Broad Iron Lines

The iron K line is intrinsically a rather narrow line. Hence, we can use broadening of the line to study the dynamics of the accretion disk. The line profiles is shaped by the effects of Doppler shifts and gravitational redshifting. Figure 8.22 demonstrates these effects at work in a schematic way. In a nonrelativistic disk, each radius of the disk produces a symmetric double-horned line profile corresponding to emission from material on both the approaching (blueshifted) and receding (redshifted) side. The inner regions of the disk, where the material is moving the fastest, produce the broadest parts of the line. Near a black hole, where the orbital velocities of the disk are mildly relativistic, special relativistic beaming enhances the blue peak of the line from each radius. Finally, the comparable influences of the transverse Doppler effect (i.e. moving clocks run slowly) and gravitational redshifting (i.e. clocks near black holes run slowly) shifts the contribution from each radius to a lower energy. Summing the line emission from all radii of the relativistic disk gives a skewed and highly broadened line profile. Turbulence in the accretion disk may also significantly affect the line profile [39].

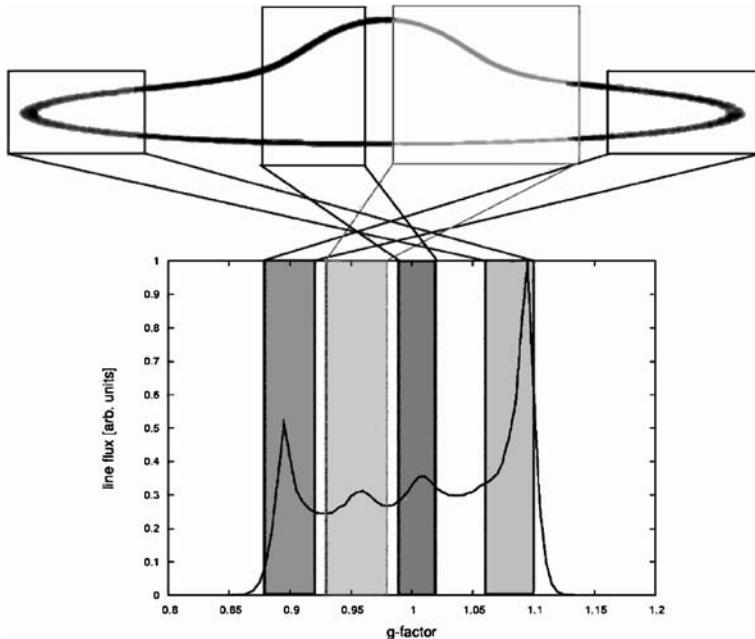


Fig. 8.22. The profile of the broad iron line emitted in the immediate vicinity of the ISCO around a black hole is caused by the interplay of Doppler and transverse Doppler shifts, relativistic beaming and gravitational redshifting. $g = v_{\text{obs}}/v_{\text{em}}$ is the Doppler factor for line emission, $g = 1$ means unshifted line. The redshifted and blueshifted wings are classical effects due to the Doppler effect from the outer disk rim, whereby the blue wing is amplified by Doppler beaming. Gravitational redshift is responsible for the large extension of the red wing. Credits: Figure provided by A. Müller (MPIE Garching)

Some fully relativistic model line profiles are plotted in Fig. 8.23, where we show the line profile from an accretion disk in orbit around a moderately rotating black hole with spin parameter $a = 0.5$ (described by the Kerr metric). The line is assumed to be emitted from an annulus of the disk extending between $6 r_g$ and $15 r_g$ from the black hole, where $r_g = GM/c^2$ is the standard gravitational radius. It is seen that the high-energy bluewards extent of the line is a strong function of the inclination of the disk. In fact, the blue extent of the line is almost entirely a function of the inclination, thereby providing a robust way to measure the inclination of the disk. On the other hand, the redward extent of the line is a sensitive function of the inner radius of the line emitting annulus. In Fig. 8.24, we show model iron lines from a Schwarzschild black hole and a rapidly rotating black hole (described by a near extremal Kerr metric). In this figure, we have made the assumption that the line emission extends down to the innermost stable orbit of the accretion disk. For these purposes, the principal difference between these two spacetime geometries is the location of the innermost stable orbit (and hence the inner edge of the line

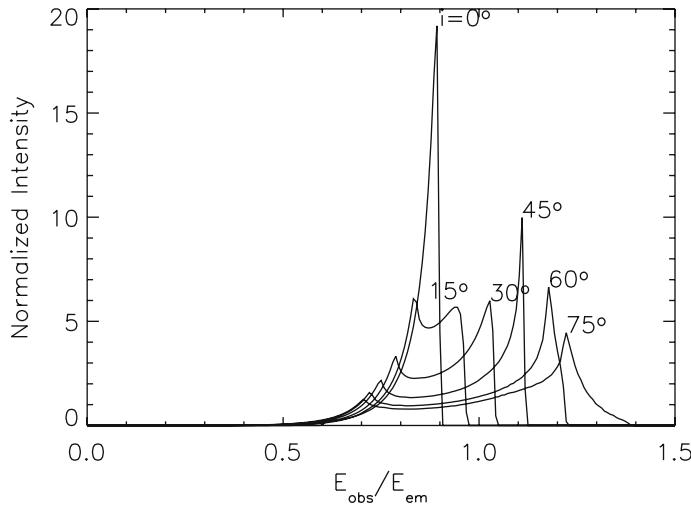


Fig. 8.23. Relativistic line profiles for the case of an accretion disk around a moderately rotating black hole $a = 0.5$. It is assumed that the fluorescing region of the disk extends from $6 r_g$ (i.e. the radius of marginal stability) to $15 r_g$ with uniform emissivity. Inclinations are shown from pole-on to edge-on view. The main effect of increasing the inclination is to broaden the line by increasing its high-energy extent

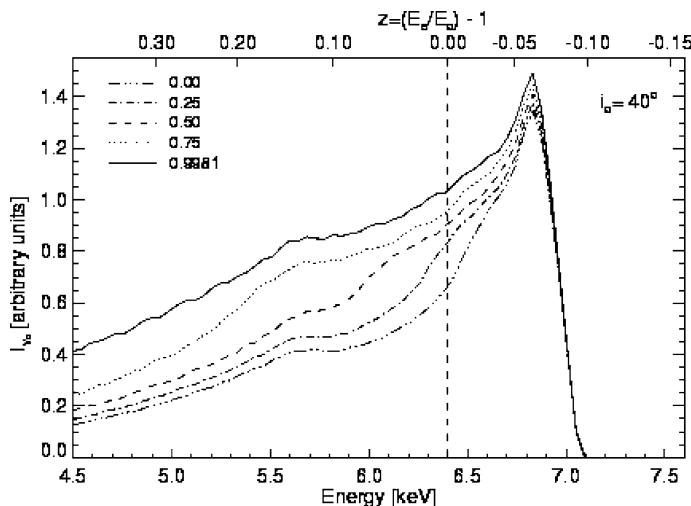


Fig. 8.24. Relativistic iron line profiles for the case of increasing Kerr parameter and a fixed inclination angle of 40 degrees. The disk is assumed to extend down to the radius of marginal stability with an emissivity index $p = 3$. Figure adapted from Müller and Camenzind [303]

emission) – this critical radius is at $6 r_g$ in the Schwarzschild case, and r_g in the extremal Kerr case.

The comparison with an observed line profile is shown in Fig. 1.1.

8.7 Dark Energy Stars

The dark energy star (often called gravastar) picture is an alternative model to the concept of a black hole, where there is an effective phase transition at or near where the event horizon is expected to form, and the interior is replaced by a de Sitter condensate. The main motivation for implementing this generalization arises from the fact that recent observations have confirmed an accelerated cosmic expansion, for which dark energy is a possible candidate.

8.7.1 Why Dark energy Stars?

Oppenheimer and Snyder [318] provided in 1939 the first insights of the gravitational collapse into a black hole. However, it was only in 1965 that marked an era of intensive research into black hole physics. Although evidence for the existence of black holes is very convincing (see next chapter), it has recently been argued that the observational data can provide very strong arguments in favor of the existence of event horizons, but cannot fundamentally prove it [23]. This scepticism has inspired new and fascinating ideas. In this line of thought, it is interesting to note that a new final state of gravitational collapse has been proposed by Mazur and Mottola [277–280]. In this model, and in the related picture developed by Laughlin et al. [117], the quantum vacuum undergoes a phase transition at or near the location where the event horizon is expected to form. The model denoted as a gravastar (*gravitational vacuum star*), consists of a compact object with an interior de Sitter condensate, governed by an equation of state given by $P = -\varrho$, matched to a shell of finite thickness with an equation of state $P = \varrho$. The latter is then matched to an exterior Schwarzschild vacuum solution. The thick shell replaces both the de Sitter and the Schwarzschild horizons, therefore, this gravastar model has no singularity at the origin and no event horizon, as its rigid surface is located at a radius slightly greater than the Schwarzschild radius. It has been argued that there is no way of distinguishing a Schwarzschild black hole from a gravastar from observational data [23]. It was also further shown by Mazur and Mottola that gravastars are thermodynamically stable. It has also been recently proposed by Chapline that this new emerging picture consisting of a compact object resembling ordinary spacetime, in which the vacuum energy is much larger than the cosmological vacuum energy, has been denoted as a **dark energy star**.

Instead of a star collapsing into a pinpoint of space with virtually infinite density, the gravastar theory proposes that as an object gravitationally collapses, space itself undergoes a phase transition preventing further collapse, being transformed into a spherical void surrounded by a form of superdense matter. In the gravastar

picture, the quantum vacuum undergoes a phase transition near the Schwarzschild radius where the event horizon would be expected to form. Externally, a gravastar appears similar to a black hole: it is visible only by the high-energy emissions it creates while consuming matter. Astronomers observe the sky for X-rays emitted by infalling matter to detect black holes, and a gravastar would produce an identical signature.

Inside a gravastar, spacetime would be totally warped by the extreme conditions there and the inner space would exert an outward force, like dark energy in the Universe. Around this vacuum region would be a shell of incredibly dense and durable matter. The phase of this matter is theorized to be similar to an extreme form of Bose–Einstein condensate in which all matter (protons, neutrons, electrons, etc.) goes into what is called a quantum state creating a kind of superatom.

In the Mazur–Mottola model, a suitable segment of the de Sitter space with an equation of state $\varrho = -P > 0$ is chosen for the interior of the compact object, while the outer region of a gravastar consists of a thin finite-thickness shell of stiff matter ($P = \varrho$) that is surrounded by a Schwarzschild vacuum ($P = 0 = \varrho$). Apart from these three layers, the Mazur–Mottola star requires two additional infinitesimally thin shells with surface densities σ_{\pm} and surface tensions that compensate the discontinuities in the pressure profile and stabilize this five-layer onion structure. Infinitesimal shells are a mathematical abstraction, physically one would like to replace the thin shells with a continuum layer of finite thickness. A clear description of this model has recently been provided by Catteon et al. [109].

In a simplified model of the Mazur–Mottola picture, Visser and Wiltshire [416] constructed a model by matching an interior solution with $P = -\varrho$ to an exterior Schwarzschild solution at a junction interface, comprising of a thin shell. The dynamic stability was then analyzed, and it was found that some physically reasonable stable equations of state for the transition layer exist. In Ref. [109, 257], a generalized class of similar gravastar models that exhibit a continuous pressure profile, without the presence of thin shells was analyzed. It was found that the presence of anisotropic pressures are unavoidable, and the TOV equation was used to place constraints on the anisotropic parameter.

8.7.2 Structure of Gravastars

Since a gravastar is a static spherical symmetric object, its structure can easily be derived by just using the EoS for vacuum energy in the interior and a normal EoS for the shell together with the TOV equation. We adopt Schwarzschild coordinates to write the line element in the form

$$ds^2 = -\exp[2\Phi(r)] dt^2 + \frac{1}{1 - 2m(r)/r} dr^2 + r^2 d\Omega^2. \quad (8.493)$$

The field equations are then those of spherical stars, except that we allow for an anisotropic pressure, $P_{\perp} \neq P_r$,

$$\frac{dm}{dr} = 4\pi\varrho r^2 \quad (8.494)$$

$$\frac{dP_r}{dr} = -\frac{(\varrho + P_r)(m + 4\pi r^3 P_r)}{r(r - 2m)} + \frac{2(P_\perp - P_r)}{r} \quad (8.495)$$

$$\frac{d\Phi}{dr} = \frac{m + 4\pi r^3 P_r}{r(r - 2m)}, \quad r < R_* \quad (8.496)$$

$$\Phi(r) = \frac{1}{2} \ln \left(1 - \frac{2M}{r} \right), \quad r \geq R_*, \quad (8.497)$$

with the boundary conditions $m(r = 0) = 0$ and $\varrho(r = 0) = \varrho_c$. Additionally to the core density ϱ_c , we have another free parameter here, namely the position of the crust r_0 , where the EoS changes from vacuum energy to the relativistic shell matter.

Physically, it is reasonable to consider static spherically symmetric geometries that satisfy:

- Inside the gravastar, $r < R_*$, the density is everywhere positive and finite.
- The central pressure is negative, $P_c < 0$, and given by $P_c = -\varrho_c$, or by the EoS of dark energy $P = w\varrho$ with $w \simeq -1$.
- The spacetime should not possess an event horizon, i.e. $2m(r) < r$ for all radii. In distinction to black holes, the Schwarzschild coordinates therefore cover the entire manifold from the center to infinity.

The pressure will roughly have a profile shown in Fig. 8.25. It naturally defines a vacuum core surrounded by a crust and some normal atmosphere. Other characteristic features of a gravastar are:

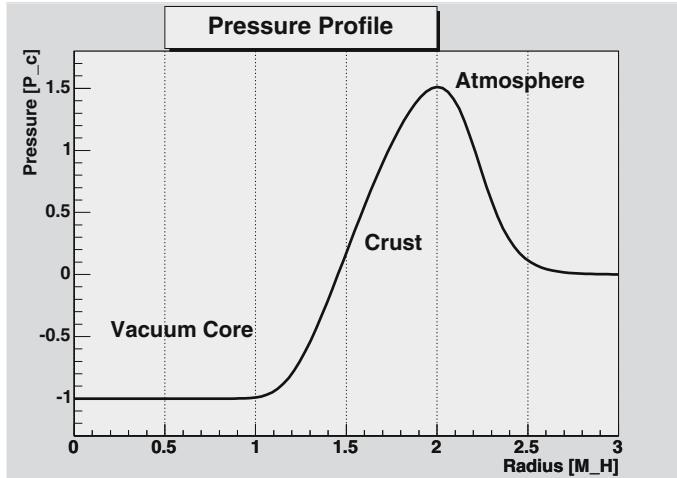


Fig. 8.25. Schematic pressure profile of a gravastar. Such a profile defines in a natural way a core, crust and atmosphere

- There should be a pressure maximum in the vicinity of the Schwarzschild radius, satisfying $P_r(r_{\max}) > 0$, and $P'_r(r_{\max}) = 0$.
- There are exactly two radii where the radial pressure vanishes, $P_r(r_0) = 0 = P_r(R)$. Clearly, $R > r_{\max}$, and R is called radius of the gravastar.
- The pressure profile P_\perp may be discontinuous, while P_r should be continuous and of the form shown in Fig. 8.25.
- The strong energy condition $\varrho + P_r + 2P_\perp \geq 0$ is definitely violated, at least near the center of the star.

The Mazur–Mottola model can then be viewed as a limiting case where $r_{\max} \rightarrow r_0$ and where an additional thin shell is placed at the surface.

8.7.3 The Necessity of an Anisotropic Crust

Assuming an isotropic EoS, $P = P_r = P_\perp$, we find for the pressure gradient at the first zero

$$\frac{dP}{dr}|_{r=r_0} = -\frac{4\pi r_0}{3} \frac{\varrho \bar{\varrho}}{1 - 2m/r_0}, \quad (8.498)$$

where we have introduced the mean density $\bar{\varrho} = m(r)/(4\pi r^3/3)$. While the left-hand side is positive, the right-hand side is negative. Therefore, we conclude that the isotropic TOV equation cannot hold at the point r_0 . The same argument holds for the location of the pressure maximum

$$\frac{dP}{dr}|_{r=r_{\max}} = -\frac{4\pi r_{\max}}{3} \frac{(\varrho + P)(\bar{\varrho} + 3P)}{1 - 2m/r_{\max}}. \quad (8.499)$$

Here, the left-hand side vanishes by definition, while the right-hand side is negative. Therefore, we also need here an anisotropic EoS. At pressure maximum, the anisotropic pressure has to compensate the inertia terms

$$2(P_\perp - P_r)|_{r=r_{\max}} = \frac{4\pi r_{\max}^2}{3} \frac{(\varrho + P_r)(\bar{\varrho} + 3P_r)}{1 - 2m/r_{\max}}. \quad (8.500)$$

Nearly all mass-energy is located in the interior vacuum energy core. The density undergoes a rapid decline in the crust region and tends asymptotically to zero. The radial component of the metric is shown in Fig. 8.26. It can be seen, that neither here, nor in the redshift factor a singularity appears. Both functions reach instead a minimum value at the position of the crust. Since the time component does also determine the redshift, a maximal redshift can be chosen by r_{\max} .

It seems that the degree of anisotropy in the crust determines the compactness $2m(r)/r$ in the crust. This would mean that the EoS of the crust is a function of the local compactness. This is certainly an unusual statement.

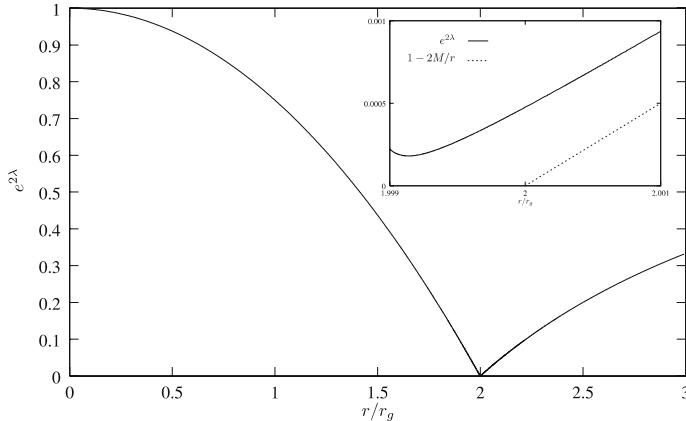


Fig. 8.26. Metric function $\exp(2\lambda)$ of a gravastar. Data provided by M. Vigelius (LSW Heidelberg)

Problems

8.1. Bound Orbits in Schwarzschild Spacetime: Investigate numerically the function $f(u)$, defined in equation (8.75) for bound orbital motion and illustrate the five different cases for the roots of $f(u)$.

Study the orbital motion of stars on the gravitational background of a Schwarzschild metric by using the closed expressions derived in Sect. 8.2 for given values of e and l .

Give closed expressions for the proper time τ and Schwarzschild time t along bound orbits. Express them in units of the Newtonian Kepler time.

Consider the special case, when $e = 0$, i.e. two roots of $f(u)$ coincide (corresponding to circular orbits). Express the proper period in terms of the Keplerian one.

8.2. Dark Energy: Find arguments in the literature, why dark energy is needed for modern cosmological models. Use the EoS of the form $P = w\rho$ to discuss the influence in the Friedman equation for the expanding Universe.

8.3. Gravastars: Use the Riemann tensors of spherically symmetric spacetimes (8.493) in Schwarzschild observers derived in Sect. 4.2.1 to discuss the tidal forces at the surface of a gravastar.

8.4. Kerr Solution for $\mu_2 + \mu_3$: The equation for the surface element of the meridional plane also results from the Einstein tensors $G_{22} = 0 = G_{33}$, since the operator $\Delta(\mu_2, \mu_3)$ cancels out, when the trace of the Ricci tensor is added. Show that they can be reduced to the following form, $\delta = \sin^2 \theta$,

$$-\frac{\mu}{\delta} (\mu_2 + \mu_3)_{,2} + \frac{r - M}{\Delta} (\mu_2 + \mu_3)_{,3} = \frac{\mu}{\varrho^2 \Delta \delta} [(r - M)(\varrho^2 + 2a^2 \delta) - 2r\Delta] \quad (8.501)$$

and

$$(r - M) \frac{\partial}{\partial r} (\mu_2 + \mu_3) + \mu \frac{\partial}{\partial \mu} (\mu_2 + \mu_3) = 2 - \frac{(r - M)^2}{\Delta} - 2 \frac{rM}{\varrho^2}. \quad (8.502)$$

These two equations are solved by means of the ansatz

$$\exp(\mu_2 + \mu_3) = \frac{\varrho^2}{\sqrt{\Delta}}. \quad (8.503)$$

The solutions for the two meridional metric functions are therefore

$$\exp(2\mu_2) = \varrho^2 / \Delta, \quad \exp(2\mu_3) = \varrho^2. \quad (8.504)$$

8.5. Komar Integrals: Use the Komar integrals,

$$M = \frac{1}{16\pi G} \oint_{\infty} dS_{\mu\nu} D^{\mu} \xi^{\nu} \quad (8.505)$$

and

$$J = \frac{1}{16\pi G} \oint_{\infty} dS_{\mu\nu} D^{\mu} m^{\nu} \quad (8.506)$$

for the total mass M and total angular momentum J of an asymptotically flat spacetime (with Killing vector ξ and m) to show that M is the mass of the gravitational source and $J = Ma$ the angular momentum of the Kerr solution.

8.6. Surface Gravity: Show that the surface gravity of the event horizon of a Kerr black hole of mass M and angular momentum J is given by

$$\kappa = \frac{\sqrt{M^4 - J^2}}{2M(M^2 + \sqrt{M^4 - J^2})}. \quad (8.507)$$

8.7. Riemann Tensors of the Kerr Geometry: By using the general expressions derived in Sect. 7.2, one can compute all components of the Riemann tensor as given by means of Bardeen observers.

8.8. Kerr Ray-Tracer: Derive the Christoffel symbols for the Kerr solution in BL coordinates. Integrate the null geodesics equations by some Runge–Kutta method and investigate the shooting of photons towards the horizon of a rotating black hole. Develop a ray-tracer for geometrically thin disks in the Kerr geometry, by applying the methods discussed in Sect. 8.6.4.

8.9. ISCO of Gravastar: Calculate the effective potential for equatorial motion in the gravitational field of a gravastar and derive the conditions for the ISCO.

9 Astrophysical Black Holes

Black holes are objects so dense that not even light can escape their gravity, and since nothing can travel faster than light, nothing can escape from inside a black hole. On the other hand, a black hole exerts the same force on something far away from it as any other object of the same mass would do. For example, if our Sun were magically crushed until it were about 3 km in size, it would become a black hole, but the Earth would remain in its same orbit.

The first type of black holes have measured masses ranging from 3 to 30 solar masses, and are believed to be formed during supernova explosions. The after-effects are observed in some X-ray binaries known as black hole candidates. On the other hand, galaxy-mass black holes are found in nearby galaxies and active galactic nuclei (AGN). These are thought to have the mass of about a few million to 10 billion solar masses. The masses of these supermassive black holes have recently been measured using various kinematic methods. X-ray observations of iron lines in the accretion disks may actually be showing the effects of such a massive black hole as well.

Additionally, there is some evidence for intermediate-mass black holes (IMBHs), those with masses of a few hundred to a few thousand times that of the Sun. These black holes may be responsible for the emission from ultraluminous X-ray sources (ULXs).

Candidates for stellar-mass black holes were identified mainly by the presence of accretion disks of the right size and speed, without the irregular flare-ups that are expected from disks around other compact objects. Stellar-mass black holes may be involved in gamma-ray bursts (GRBs); short duration GRBs are believed to be caused by colliding neutron stars, which form a black hole on merging. Observations of long GRBs in association with supernovae suggest that long GRBs are caused by collapsars – a massive star whose core collapses to form a black hole, drawing in the surrounding material. Therefore, a GRB could possibly signal the birth of a new black hole.

From observations in the 1980s of motions of stars around the Galactic center, it is now believed that such supermassive black holes exist in the center of most galaxies, including our own Milky Way. Sagittarius A* is now generally agreed to be the location of a supermassive black hole at the center of the Milky Way galaxy. The orbits of stars within a few AU of Sagittarius A* rule out any object other than a black hole at the center of the Milky Way assuming the current standard laws of physics are correct.

The current picture is that all galaxies may have a supermassive black hole in their center and that this black hole accretes gas and dust in the middle of the galaxies generating huge amounts of radiation – until all the nearby gas has been swallowed or dispersed and the process shuts off. This picture also nicely explains why there are no nearby quasars.

9.1 Classes of Astrophysical Black Holes

Black holes in the Universe are naturally subdivided into five classes according to their masses (see Fig. 1.9)

- **Primordial black holes** with masses in the range between an Earth mass and a solar mass. They could result from the early Universe, but their formation

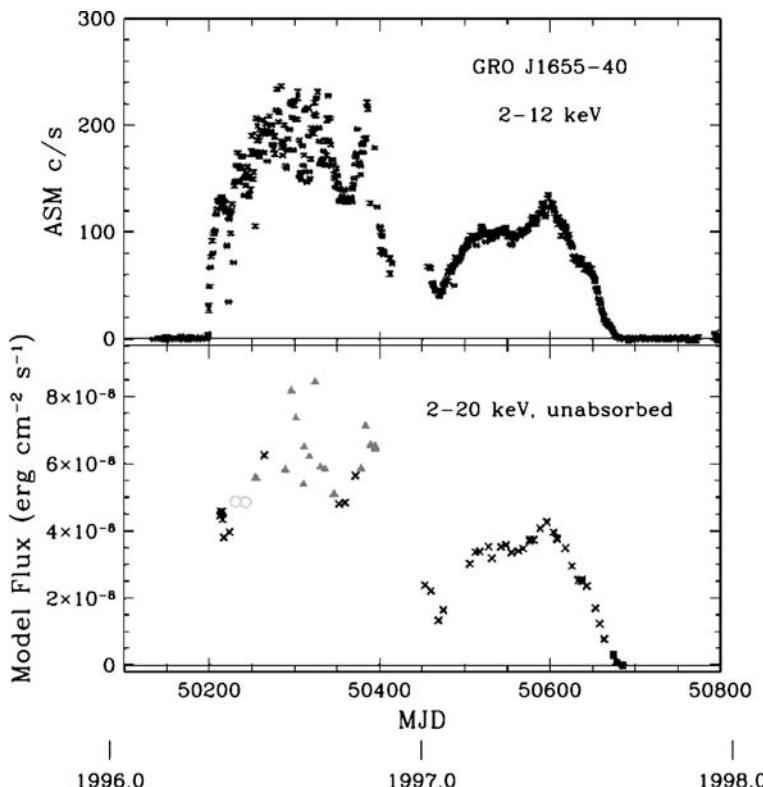


Fig. 9.1. Evolution of X-ray states during the nova outburst of GRO J1655–40 detected with ASM on RXTE. *Top panel* shows the ASM light-curve. *Bottom panel* shows fluxes from pointed observations. The symbol type denotes the X-ray state: *crosses* mark thermal states (black-body spectra), *squares* hard states, *triangles* steep power laws, and *circles* intermediate states. Data from Remillard [340]

is completely unclear. Primordial black holes could provide a unique probe of the early Universe, gravitational collapse, high-energy physics and quantum gravity [104].

- **Stellar black holes** with masses $2.5 M_{\odot} \leq M_H \leq 50 M_{\odot}$. These objects are formed in evolution of massive stars and accretion onto neutron stars. As of 2005, about 20 stellar black holes have been found in X-ray binary systems. No stellar black hole has so far been detected as companion of a radio pulsar.
- **Intermediate mass black holes** with masses $50 M_{\odot} < M_H < 10^5 M_{\odot}$. It is presently unclear whether objects are formed in this intermediate mass range. Ultraluminous X-ray sources (ULXs) are a distinctive class of objects, because they radiate 10 to 1000 times more X-ray power than neutron stars and stellar mass black holes.
- **High-redshift black holes** with masses $10^5 M_{\odot} < M_H < 10^6 M_{\odot}$. They could be formed at redshifts of $\simeq 26$ by direct collapse of supermassive clouds in massive dark halos.
- **Supermassive black holes** with masses $10^6 M_{\odot} \leq M_H \leq 10^{10} M_{\odot}$. These objects are formed in the center of spheroidal galaxies, and they have been proven to exist in almost all galaxies. In large galaxy surveys (2dF and SDSS), more than 100,000 supermassive black holes have been detected in quasars.

Candidates for supermassive black holes were first provided by the active galactic nuclei and quasars, discovered by radioastronomers in the 1960s. The efficient conversion of mass into energy by friction in the accretion disk of a black hole seems to be the only explanation for the copious amounts of energy generated by such objects. Indeed the introduction of this theory in the 1970s removed a major objection to the belief that quasars were distant galaxies – namely, that no physical mechanism could generate that much energy.

9.2 Measuring Black Hole Masses

The most accurate mass measurements in astrophysics are obtained via dynamical methods.

9.2.1 BHs in X-Ray Binaries

In the case of BHs in X-ray binary systems (XRBs), it is a standard method to measure the orbital period P_b and the maximum line-of-sight Doppler velocity $K_c = v \sin i$ of the companion star, where i is the inclination angle of the binary orbit. From this information, one calculates the mass function

$$f(M_H, i) = \frac{K_c^3 P_b}{2\pi G} = \frac{M_H \sin^3 i}{(1 + M_H/M_c)^2}. \quad (9.1)$$

M_H is the mass of the BH candidate, and M_c the mass of the companion star. Often, one has an independent estimate for the inclination angle i . In addition, one requires

an independent estimate for the mass M_c , e.g. from spectral type (see Table 9.1). In Cyg X-1, M_c is large (the companion is a blue supergiant) and seriously affects the estimate for M_H . In low-mass type binary systems, such as A0620–00 and other X-ray novae (see Table 9.1), the companion mass is often less than 10 percent of M_H .

By combining measurements of K_c and P_b with estimates for i and M_c , the masses of the compact X-ray emitting stars in a number of XRBs have been measured [283]. The confirmed stellar BH candidates are listed in Table 9.1. Most of the sources are LMXB transients.

Cygnus X-1 is an X-ray binary that was one of the first X-ray sources discovered when it was detected in 1962. The visible object HDE226868 is a ninth magnitude blue O9 supergiant star whose radial velocity curve shows an orbital period of 5.6 days. The fact that the object is a strong X-ray emitter and that the optical and X-ray emission varies on very short time-scales (as short as one one-thousandth of a second) suggests that the companion might be a black hole. Analysis of the radial velocity variation of the primary under the assumption that it is a normal star suggests that the mass of the companion is about 10 solar masses.

LMC X-3 is a powerful source of X-rays located in the Large Magellanic Cloud. The X-ray source is associated with a binary system with an orbital period of 1.7 days. The visible component is a main sequence B3 star whose shape has been severely

Table 9.1. Confirmed stellar black hole candidates. All masses are given in solar-mass units. Data adapted from McClintock and Remillard [283], and Casares [108]. μ Q indicates that these sources are pronounced microquasars

Object	P_b [d]	$\frac{M_H^3 \sin^3 i}{(M_H + M_c)^2}$	Donor	M_H	Classification
GRS 1915+105	33.5	9.5 ± 3	K/M III	14 ± 4	LMXB Transient μ Q
V404 Cyg	6.470	6.08 ± 0.06	K0 IV	12 ± 2	LMXB Transient
Cyg X-1	5.600	0.244 ± 0.005	O9.7Iab	14 ± 4	HMXB Persistent μ Q
LMC X-1	4.229	0.14 ± 0.05	O7 III	> 4	HMXB Persistent
LMC X-3	1.704	2.3 ± 0.3	B3 V	7.6 ± 1.3	HMXB Persistent
GRO J1655–40	2.620	2.73 ± 0.09	F3 IV	6.3 ± 0.3	LMXB Transient μ Q
XTE J1819–254	2.816	3.13 ± 0.13	B9 III	7.1 ± 0.3	LMXB Transient
GX 399–4	1.754	5.8 ± 0.5	–	–	LMXB Transient μ Q
XTE J1550–564	1.542	6.86 ± 0.71	G8 IV	9.6 ± 1.2	LMXB Transient
4U 1543–47	1.125	0.25 ± 0.01	A2 V	9.4 ± 1.0	LMXB Transient
H 1705–250	0.520	4.86 ± 0.13	K3 V	6 ± 2	LMXB Transient
GS 1124–168	0.433	3.01 ± 0.15	K3 V	7.0 ± 0.6	LMXB Transient
XTE 1859+226	0.382	7.4 ± 1.1	–	–	LMXB Transient
GS 2000+25	0.345	5.01 ± 0.12	K3 V	7.5 ± 0.3	LMXB Transient
A 0620–00	0.325	2.72 ± 0.06	K4 V	11 ± 2	LMXB Transient
XTE J1650–500	0.321	2.73 ± 0.56	K4 V	–	LMXB Transient
GRS 1009–45	0.283	3.17 ± 0.12	K7 V	5.2 ± 0.6	LMXB Transient
GRO J0422+32	0.212	1.19 ± 0.02	M2 V	4 ± 1	LMXB Transient
XTE J1118+480	0.171	6.3 ± 0.2	K5 V	6.8 ± 0.4	LMXB Transient

distorted by the gravitational field of its companion. Although not unambiguous, the mass of the compact object is estimated to be at least four solar masses and more likely is considerably higher, making this one of the best black hole candidates (Fig. 9.4).

In recent years, XRBs have been found which show jets in the radio regime, similar to extragalactic jet sources. For this reason, they have been called *microquasars* [297]. Previously identified as a source of variable X-ray emission, the source known as GRS 1915+105 recently and suddenly showed a rapid increase in its radio flux (Fig. 9.2). Radio astronomers mapping the source have shown that two clumps of high-energy particles are being ejected simultaneously from the source at velocities close to the speed of light. The object is probably a collapsed star, a neutron star or black hole formed in a supernova explosion, in orbit with a more normal companion. Material from the companion spirals toward the collapsed remnant, is heated, and thus emits X-rays. Jets of subatomic particles are ejected in opposite directions away from the compact object; high-velocity electrons emit synchrotron radiation, detected at radio wavelengths.

By observing the radio emission over several months, Mirabel and Rodriguez [296] were able to see that the radio emitting plasma blobs were moving. Their apparent motion (change of position with time) implies that they are travelling with velocities greater than the speed of light – an impossibility according to Einstein’s theory – and are thus superluminal. However, the apparent motion can also be understood as the result of an illusion caused by the combination of their high velocity (close to but

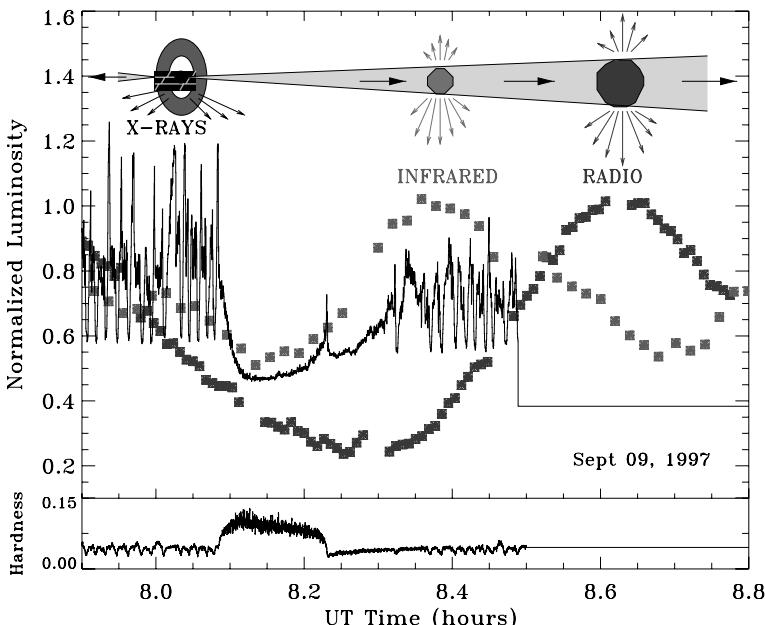


Fig. 9.2. Light-curves of the microquasar GRO 1915+105. Figure adapted from [297]

less than the speed of light) and the orientation of the jets along the direction close to the line-of-sight.

Microquasars are X-ray binary stars which exhibit jets of relativistic matter ejected perpendicularly to both sides of an accretion disk (Fig. 9.3). The jets contain relativistic electrons that produce the synchrotron radiation detected at radio wavelengths. GRO J1655–40 is the second so-called “microquasar” discovered in our Galaxy. Microquasars are black holes of about the same mass as a star. They behave as scaled-down versions of much more massive black holes that are at the cores of extremely active galaxies, called quasars (Fig. 9.3). Its recent outburst detected with the ASM monitor onboard of RXTE shows a complicated time structure (Fig. 9.1). This surprising result showed that the supermassive black holes at the centers of galaxies – black holes millions of times more

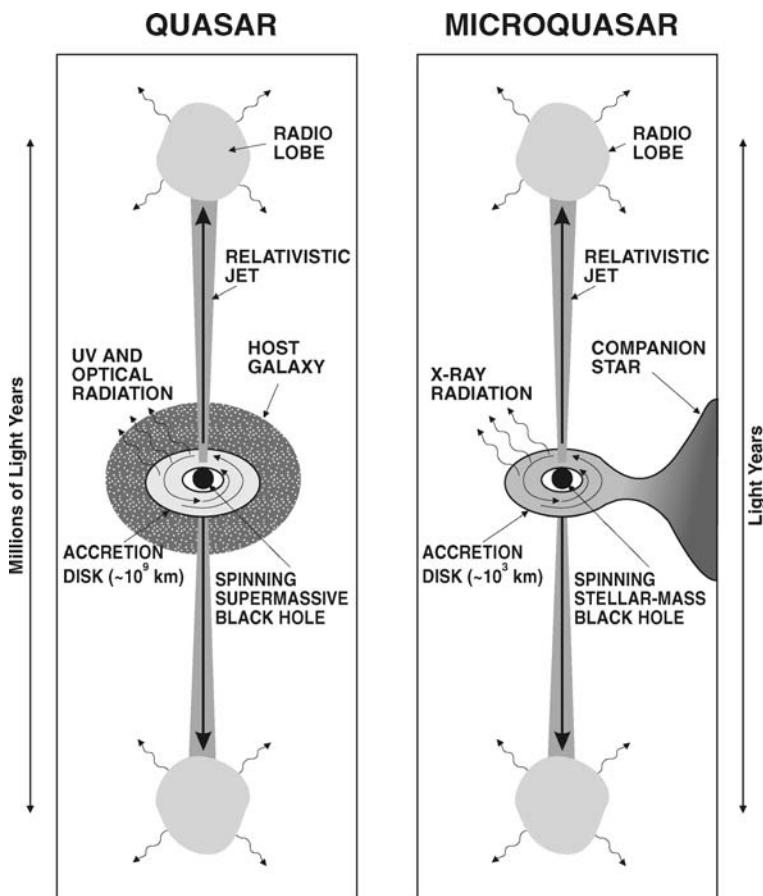


Fig. 9.3. Microquasars are scaled-down versions of quasars. Figure adapted from Mirabel and Rodriguez [297]

massive than the Sun – have smaller counterparts capable of producing similar jet ejections. GRS 1915+105 is thought to be a double star system, in which one of the components is a black hole about 14 times the mass of the Sun. The more massive object is pulling material from its stellar companion. The material circles the massive object in an accretion disk before being pulled into it. Friction in the accretion disk creates temperatures hot enough that the material emits X-rays, and magnetic processes are believed to accelerate the material in the jets.

The small size of microquasars makes them especially easy to monitor in time. Considering that the characteristic time of matter accretion onto the compact object is proportional to its mass, the variations observed in a microquasar in time-scales of minutes (easy to follow by a human observer) correspond to similar phenomena that would take thousands of years in a quasar with, say, one billion solar masses.

Using a new technique astronomers have recently found 10 apparent black holes near the center of the Andromeda Galaxy, the nearest large spiral galaxy to our own. The search method might be employed to uncover more black holes in our Milky Way and in other, more distant galaxies. Andromeda is 2.5 million lightyears away. The newfound black hole candidates are of the stellar variety, meaning they are several times the mass of the Sun and are the collapsed remains of dead stars. Each has a companion object, an orbiting normal star that feeds material to the black hole.

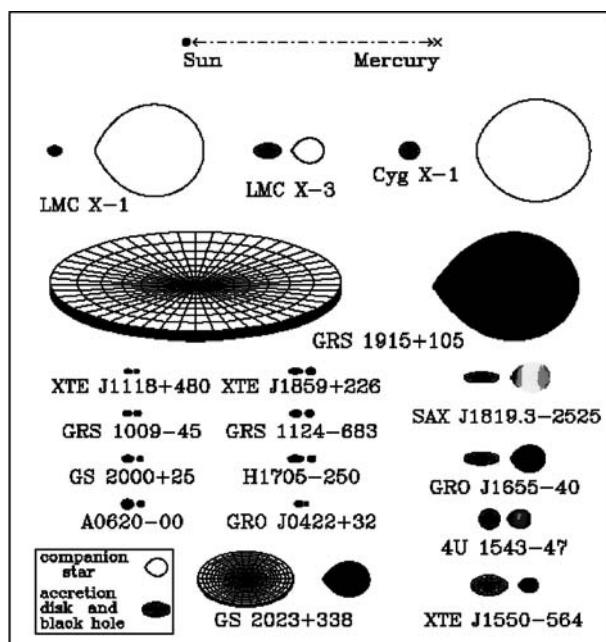


Fig. 9.4. This is a schematic diagram of 17 dynamically confirmed black hole binaries. Fourteen have low mass companions (i.e. stars with masses less than about 3 solar masses), and the three on the top have high mass companions. Credit: with permission by Orosz

9.2.2 Intermediate-Mass Black Holes

Mysterious, powerful X-ray sources found in nearby galaxies may represent a new class of objects, according to data from NASA's Chandra X-ray Observatory (see Miller and Colbert [295]). These sources, which are not as hot as typical neutron star or black hole X-ray sources, could be a large new population of black holes with masses several hundred times that of the Sun. Until a few years ago, astronomers only knew of two sizes of black holes: stellar black holes, with masses about 10 times the Sun, and supermassive black holes located at the centers of galaxies, with masses ranging from millions to billions times the Sun. Recent evidence suggests a class of *intermediate-mass black holes* (IMBHs) may also exist.

However, it is not clear how such a black hole would form. On the one hand, they are too massive to be formed by the collapse of a single star, which is how the stellar black holes are thought to form. On the other hand, their environments lack the extreme conditions, i.e. high density and velocities observed at the centers of galaxies, which seemingly lead to the formation of supermassive black holes. There are two popular formation scenarios for IMBHs. The first is the merging of stellar mass black holes and other compact objects by means of gravitational radiation. The second one is the runaway collision of massive stars in dense stellar clusters and the collapse of the collision product into an IMBH.

9.2.3 Supermassive Black Holes in Nearby Galaxies

Black holes with masses of a million to a few billion times the mass of the Sun are believed to be the engines that power nuclear activity in galaxies. Active nuclei range from faint, compact radio sources like that in M31 to quasars like 3C 273 that are brighter than the whole galaxy in which they live (see Fig. 10.25). Some nuclei fire jets of energetic particles millions of lightyears into space. Almost all astronomers believe that this enormous outpouring of energy comes from the death throes of stars and gas that are falling into the central black hole. This is a very successful explanation of the observations, but until recently, it was seriously incomplete: we had no direct evidence that supermassive black holes exist.

For the past 20 years, astronomers have looked for supermassive black holes by measuring rotation and random velocities of stars and gas near galactic centers. If the velocities are large enough, as in the Sombrero Galaxy, then they imply more mass than we see in stars. The most probable explanation is a black hole. About 50 have been found as of 2005. Their masses are in the range expected for nuclear engines, and their numbers are consistent with predictions based on the energy output of quasars (see Table 9.2).

A giant black hole in a galactic nucleus exerts a powerful gravitational force on nearby gas and stars, causing them to move at high speeds. This is hard to see in quasars, because they are far away and because the dazzling light of the active nucleus swamps the light from the host galaxy. In a radio galaxy with a fainter nucleus, the stars and gas are more visible. The giant elliptical galaxy Messier 87, one of the two brightest objects in the Virgo cluster of galaxies, is a radio galaxy

Table 9.2. Black holes masses in nearby galaxies, data adapted from [189] and references therein

Galaxy	Type	σ_* [km/s]	L_{Bulge} [L_\odot]	M_{Bulge} [M_\odot]	M_H [M_\odot]
GC	SBc	75	–	1.1×10^{10}	$(3.7 \pm 1.5) \times 10^6$
M31	Sb	160	7.3×10^9	3.7×10^{10}	$(4.5 \pm 3.0) \times 10^7$
M32	E2	75	3.8×10^8	8.0×10^8	$(2.5 \pm 0.5) \times 10^6$
M87	E0	375	2.0×10^{11}	6.0×10^{11}	$(3.0 \pm 1.0) \times 10^9$
N1068	Sb	151	1.5×10^{11}	2.3×10^{11}	$(1.4 \pm 1.0) \times 10^7$
N3379	E1	206	1.7×10^{10}	6.8×10^{10}	$(1.0 \pm 0.5) \times 10^8$
N4374	E1	296	6.0×10^{10}	3.6×10^{11}	$(4.3 \pm 2.0) \times 10^8$
N4261	E2	315	4.5×10^{10}	3.6×10^{10}	$(5.2 \pm 1.1) \times 10^8$
N6251	E2	290	9.3×10^{10}	5.6×10^{11}	$(5.3 \pm 3.0) \times 10^8$
N7052	E4	266	8.3×10^{10}	2.9×10^{11}	$(3.3 \pm 2.0) \times 10^8$
N4742	E4	90	6.2×10^9	6.2×10^9	$(1.4 \pm 0.4) \times 10^7$
N821	E4	209	2.9×10^9	1.3×10^{11}	$(3.7 \pm 1.5) \times 10^7$
IC459	E3	323	6.9×10^{10}	2.9×10^{11}	$(2.5 \pm 0.5) \times 10^9$
N1023	SB0	205	1.2×10^{10}	6.9×10^{10}	$(4.4 \pm 0.4) \times 10^7$
N2778	E2	175	1.2×10^{10}	7.6×10^{10}	$(1.4 \pm 1.2) \times 10^7$
N3115	S0	230	1.7×10^{10}	1.2×10^{11}	$(1.0 \pm 0.8) \times 10^9$
N3245	S0	205	1.7×10^{10}	6.8×10^{10}	$(2.1 \pm 0.8) \times 10^8$
N3377	E5	145	6.4×10^9	1.6×10^{10}	$(1.0 \pm 0.9) \times 10^8$
N3384	S0	143	7.1×10^9	2.0×10^{10}	$(1.6 \pm 0.2) \times 10^7$
N3608	E2	182	1.9×10^{10}	9.7×10^{10}	$(1.9 \pm 0.8) \times 10^8$
N4291	E2	242	1.9×10^{10}	1.3×10^{11}	$(31. \pm 1.2) \times 10^8$
N4342	S0	225	1.9×10^9	1.2×10^{10}	$(3.0 \pm 1.4) \times 10^8$
N4473	E5	190	1.8×10^{10}	9.2×10^{10}	$(1.1 \pm 0.8) \times 10^8$
N4564	E3	162	8.1×10^9	4.4×10^{10}	$(5.6 \pm 0.5) \times 10^7$
N4594	Sa	240	4.4×10^{10}	2.7×10^{11}	$(1.0 \pm 0.8) \times 10^9$
N4649	E1	376	6.1×10^{10}	4.9×10^{11}	$(2.0 \pm 0.8) \times 10^9$
N4697	E4	177	2.3×10^{10}	1.1×10^{11}	$(1.7 \pm 0.2) \times 10^8$
N5845	E3	234	6.7×10^9	3.7×10^{10}	$(2.4 \pm 1.0) \times 10^8$
N7332	S0	122	7.9×10^9	1.5×10^{10}	$(1.3 \pm 0.6) \times 10^7$
N7457	S0	67	2.1×10^9	7.0×10^9	$(3.5 \pm 1.2) \times 10^6$

with a bright jet emerging from its nucleus (Fig. 10.22). It has long been thought to contain a black hole. Recent observations of Messier 87 with the Hubble Space Telescope (HST) revealed a disk of gas 500 lightyears in diameter, whose orbital speeds imply a central mass of three billion solar masses. The ratio of this mass to the central light output is more than 100 times the solar value. No normal population of stars has such a high mass-to-light ratio. This is consistent with the presence of a black hole, but it does not rule out some other concentration of underluminous matter.

To get a sense of how stellar dynamics are used to measure BH masses, consider the idealized case of spherical symmetry. Then the first velocity moment of the

collisionless Boltzman equation will give the mass of the system as a function of radius (called the Jeans equation)

$$M(r) = \frac{V^2 r}{G} + \frac{\sigma_r^2 r}{G} \left[-\frac{d \ln \varrho_*}{d \ln r} - \frac{d \ln \sigma_r^2}{d \ln r} - \left(1 - \frac{\sigma_\theta^2}{\sigma_r^2} \right) - \left(1 - \frac{\sigma_\phi^2}{\sigma_r^2} \right) \right]. \quad (9.2)$$

The rotation velocity V , the velocity dispersions σ_i^2 , and the density ϱ_* are, in principle, all measurable, and so a mass can be found down to the resolution limit of the instrument. There are, however, several immediate problems. First is that we observe these quantities projected on the sky, and de-projecting these quantities to compare with observations is a difficult process. Second, as mentioned above, most galaxies are at least axisymmetric; spherical symmetry is unlikely. Anisotropy must be explored, usually for example by using the “maximum entropy” method. The state-of-the-art in using stellar dynamical analysis to get BH masses is thus to do the following. High quality observations of spectra are required (HST) and the line-of-sight velocity dispersion (LOSVD) must be extracted from the data. The LOSVD is fitted with three-integral models of the galaxy, using an orbit based approach, known as Schwarzschild’s method. Maximum entropy models are used to account for axisymmetry, and the potential is derived from the surface brightness profile.

A more compelling argument is possible in the Seyfert galaxy NGC 4258, where orbiting gas in the nucleus emits microwave maser emission from water molecules (Fig. 9.5). The location and velocity of this gas can be mapped with amazing precision by making coordinated observations with radio telescopes separated by large distances. The angular resolution given by this technique is 100 times better than that of HST. The measurements imply that 40 million solar masses lie within half a lightyear of the center. Could this material be a cluster of dark stars? The answer is no.

There are two possibilities, failed stars or dead stars. Failed stars are ones that are too low in mass; their insides never get hot enough to ignite the nuclear reactions that power stars. They are called brown dwarfs. But brown dwarfs are light – less than 0.08 solar masses – so there would have to be many of them to explain the dark mass in NGC 4258. Then they would have to live very close together. As a result, most of them would collide with other brown dwarfs. Stars that collide generally stick together. But if two brown dwarfs of almost 0.08 solar masses merge, they become a luminous star, and then the dark cluster would light up. The other alternative is dead stars, that is, white dwarf stars, neutron stars, or stellar-mass black holes. But these are more massive than brown dwarfs, so there would be fewer of them. The gravitational evolution of clusters of stars is well understood: individual stars get ejected from the cluster, the remaining cluster contracts, and the evolution speeds up. Calculations show that a cluster of dead stars in NGC 4258 would evaporate completely in about 100 million years. From a cosmic perspective, this is almost no time at all. It is much less than the age of the galaxy. So the most astrophysically plausible alternatives to a black hole can be excluded. It is difficult to escape the conclusion that NGC 4258 contains a supermassive black hole.

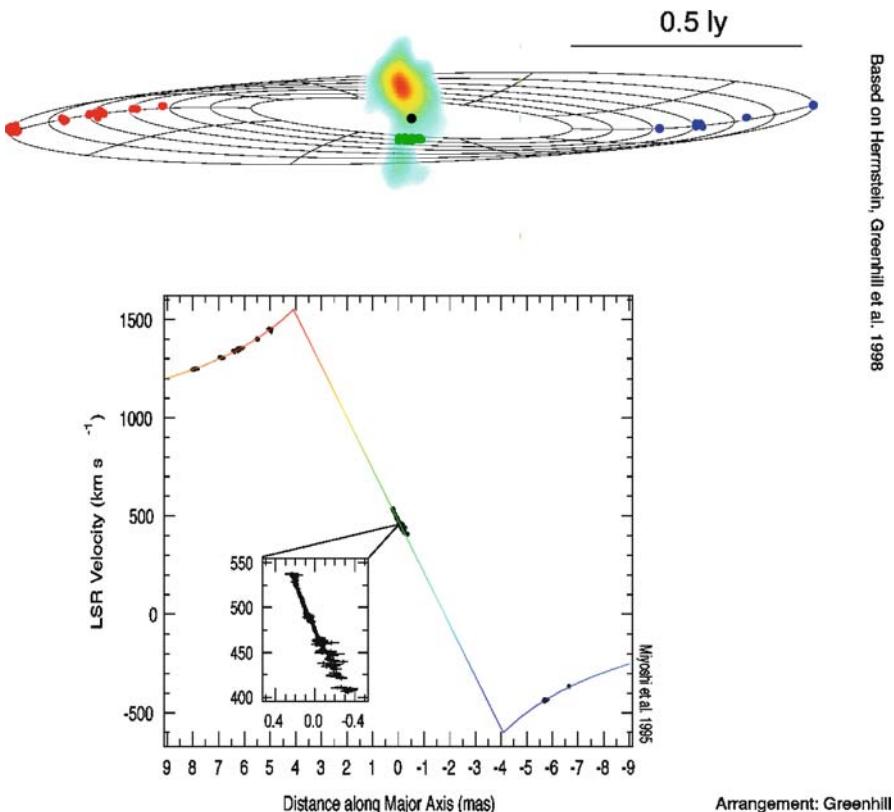


Fig. 9.5. *Upper panel:* The spots show the relative positions of maser emitting clouds in the gas disk at the center of NGC 4258, determined by radio interferometry. The black dot is the position of the central supermassive BH. *Bottom panel:* The dots show the line-of-sight velocities of the maser spots as a function of distance from the BH along the major axis of the disk. These dots follow perfectly the velocities of Keplerian motion. The mass inferred from these measurements is $M_H = (3.5 \pm 0.1) \times 10^7 M_\odot$. Figures adapted from Greenhill et al. [183]; Herrnstein et al. [201]

NGC 3115 is a nearby S0 galaxy, an intermediate type between spirals and ellipticals in Edwin Hubble's classification scheme. Spectroscopic observations made in 1992 with the Canada–France–Hawaii telescope revealed rapid rotation from an edge-on disk and large random velocities in the surrounding bulge stars. These implied a black hole mass of one billion suns. In 1996, observations with HST confirmed this result and showed a velocity dispersion as high as 600 km/s in the nucleus. This was the highest measured in any galactic center. NGC 3115 contains a tiny nuclear star cluster, exactly the sort of density cusp that astronomers have long expected to find around black holes. This nucleus would fly apart in a few tens of thousands of years unless there is a dark mass equivalent to a billion suns holding it together.

Based on Herrnstein, Greenhill et al. 1998

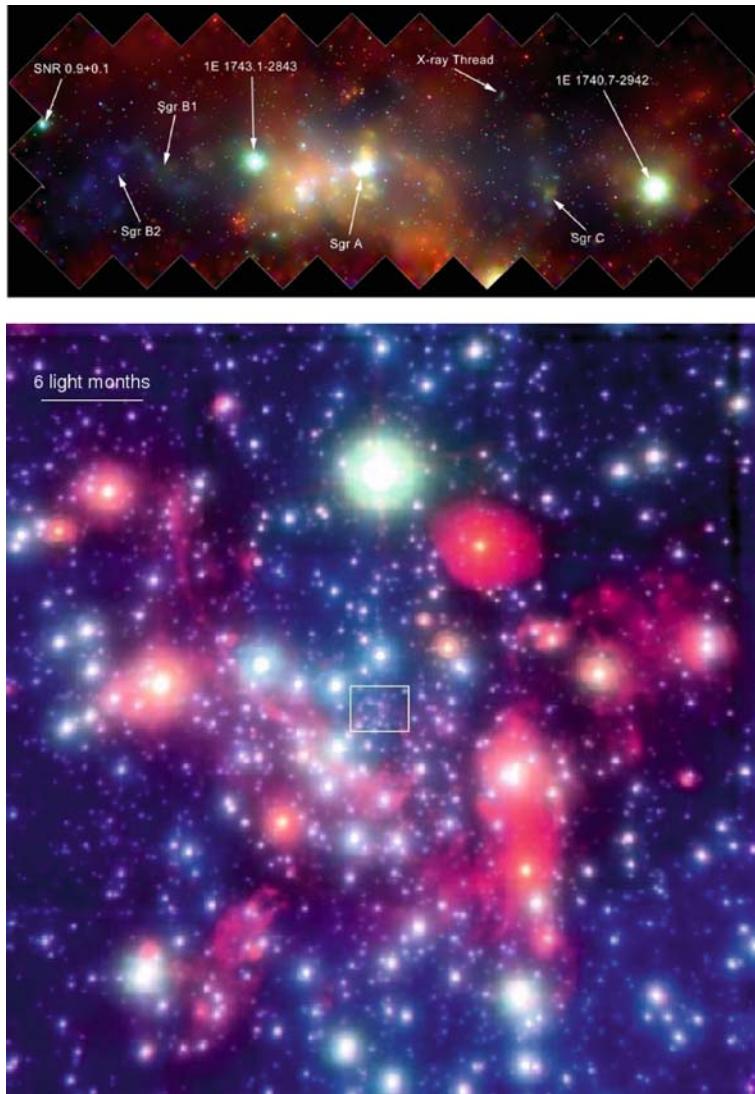


Fig. 9.6. The Galactic center and its black hole. *Top:* This 400×900 lightyear mosaic of several Chandra images of the central region of our Milky Way galaxy reveals hundreds of white dwarf stars, neutron stars, and black holes bathed in an incandescent fog of multimillion-degree gas (Credit: image credit Chandra (NASA)). The supermassive black hole at the center of the Galaxy is located inside the bright white patch in the center of the image. Credits: Chandra X-Ray Observatory. *Bottom:* A HKL color composite of the Galactic center region. The central black hole is located in the center of the box. This image has kindly been provided by A. Eckart

The BH in the Galactic Center

Is there a black hole in our own galaxy, the Milky Way? The center of our Galaxy is only 25,000 lightyears away, but its visible light is completely absorbed by intervening dust. Fortunately, infrared light penetrates the dust (Fig. 9.6). Early infrared measurements suggested that we have a black hole of about a million solar masses. It is coincident with a weak radio source called Sgr A* that is exceptionally tiny – radio observations show that it is about the size of Saturn’s orbit around the Sun. In a dramatic breakthrough, two groups of astronomers led by Reinhard Genzel in Munich, Germany, and Andrea Ghez at UCLA, have measured the motions of individual stars near the Galactic center as projected on the plane of the sky (Fig. 9.7). They used a technique called speckle imaging to reduce the blurring effect of the Earth’s atmosphere. The velocity dispersion increases to 400 km/s at a distance of

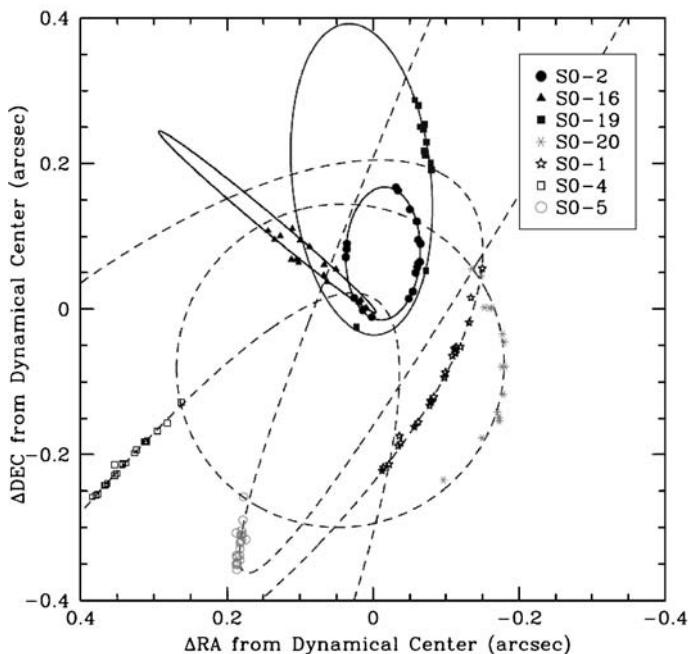


Fig. 9.7. Astrometric positions and orbital fits for eight stars, within the central 0.8×0.8 arcsec of the Galaxy, that show significant deviation from linear motion. The proper motion measurements were obtained at the Keck telescopes between 1995 and 2003, have uncertainties that are comparable to the size of the points, and are plotted in the reference frame in which the central dark mass is at rest. Overlaid are the best fitting simultaneous orbital solutions, which assume that all the stars are orbiting the same central point mass. The orbital solutions for the three stars that constrain the properties of the central dark object are delineated by solid lines and the joint orbital solutions for the remaining stars are shown with dashed lines. These orbits provide the best evidence yet for a supermassive black hole, which has a mass of 3.7 million times the mass of the Sun. Figure adapted from Ghez et al. [173]

0.03 lightyears from Sgr A*. Here stars have such small orbits that they revolve around the Galactic center in a few decades. We can look forward to seeing the Galactic center rotate once in our lifetimes! Motions in the plane of the sky have been measured for these stars for several years, and such velocity measurements have strengthened the case for a 3.6 million solar mass black hole at the center of our Galaxy. The new observations for the first time allow the measurement of accelerations (and not just velocities) for three of the stars. The acceleration vectors intersect, to within errors, at Sgr A*, confirming the identification of the black hole with the radio source. The rapid motions show that there is a mass of three million solar masses centered on Sgr A* (Fig. 9.8). Notice how the mass enclosed inside a particular distance from Sgr A* stops dropping toward the center at a distance of about 1 pc. This means that the mass in stars inside three lightyears has become negligible compared to the dark mass at the center. As in NGC 4258, the implied density of matter is too high to allow a cluster of dark stars or stellar remnants. Because of this work, the most compelling example of a supermassive black hole is in our own Galaxy.

The spectrum of Sgr A* in the radio can be described by a power law with index 0.3 between gigahertz and millimeter wavelengths. The most interesting feature is the existence of a bump in the submillimeter regime, which is probably polarized (Fig. 9.9). This submillimeter excess in the spectrum may be the first evidence for gas falling towards the black hole on scales of only a few gravitational radii. Since Sgr

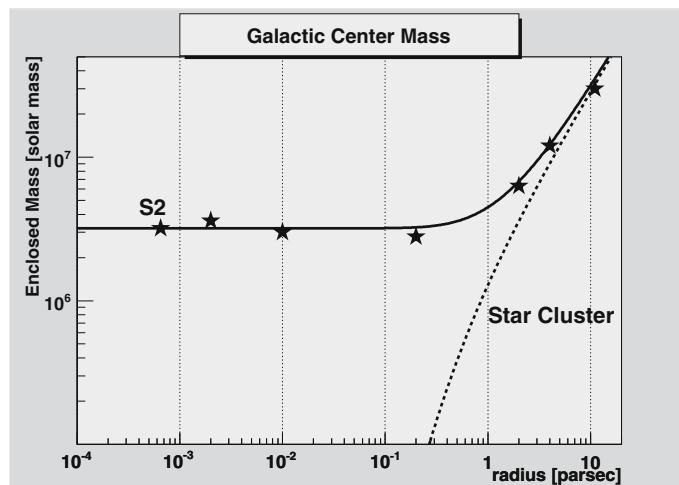


Fig. 9.8. Central mass distribution in our Galaxy implied by the observed velocities measured on the parsec and subparsec scale. The *solid curve* represents the stars plus a point mass of 3.2 million suns. The *dashed curve* gives the contribution from the star cluster on the parsec-scale which has a central density $\rho_c = 3 \times 10^6 M_\odot \text{ pc}^{-3}$ and a core radius $r_c = 0.3 \text{ pc}$. A corresponding star cluster to explain the data point S2 would require a central density $\rho_c = 2 \times 10^{17} M_\odot \text{ pc}^{-3}$ and a core radius of only 0.0002 pc. This is excluded on astrophysical grounds

A^* is embedded into a nuclear star cluster including He stars with high mass losses, Bondi accretion would result from the innermost region with low angular momentum (Melia et al. [289]). A disk is however formed at a distance of a few gravitational radii. Coker and Melia [121] concluded that the accreted specific angular momentum $j = \lambda R_{SC}$ has an average equilibrium value for $\lambda \simeq 30$ or less. As a result of this, the flow would circularize at a radius of about 5–50 gravitational radii, but no extended disk would form in the Galactic center. The bump in the spectrum would result as thermal cyclotron emission from hot relativistic electrons in the innermost part of the disk with magnetic fields of the order of 100 gauss. The main contribution results from cyclotron emission with peak frequency at about 100 GHz and strong self-absorption below this frequency (due the extreme compactness of the emission region). A second peak results from Comptonization of the cyclotron emission by the hot electrons with

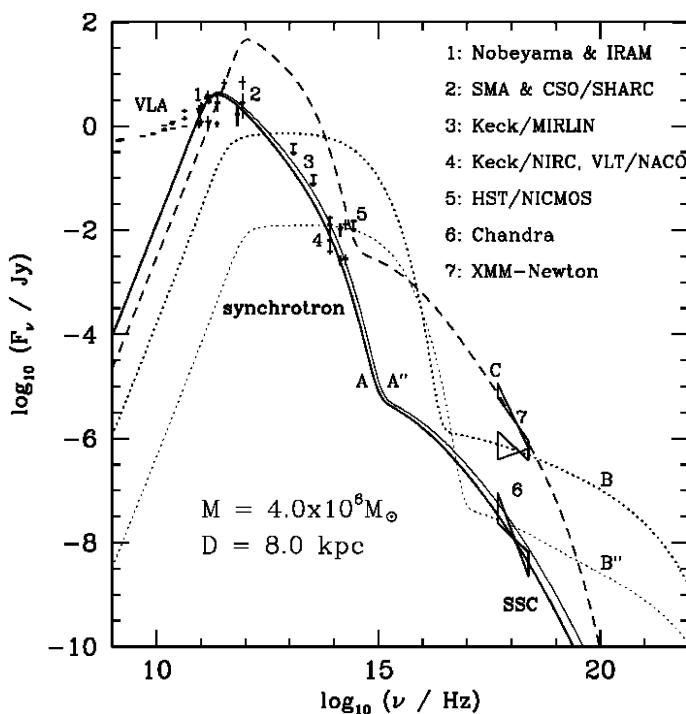


Fig. 9.9. Overall spectrum emitted by hot electrons in the inner disk around Sgr A^* . The data are compiled by Liu et al. [256] for different epochs, the instruments are indicated. Continuous lines are some model fits. Acceleration of electrons by plasma wave turbulence in hot gases near the black hole's event horizon can account both for Sagittarius A^* 's millimeter and shorter wavelengths emission in the quiescent state, and for the infrared and X-ray flares, induced either via an enhancement of the mass accretion rate onto the black hole or by a reorganization of the magnetic field coupled to the accretion gas. Figure adapted from Liu et al. [256]

a temperature of $\simeq 100$ MeV. The Chandra spectrum is formed by this part of the emission. The disk around Sgr A* would emit asymmetrically (Fig. 9.10).

A swarm of 10,000 or more black holes may be orbiting the Milky Way's supermassive black hole, according to new results from NASA's Chandra X-ray Observatory. This would represent the highest concentration of black holes anywhere in the Galaxy [306]. These images (see Fig. 9.6) are part of an ongoing Chandra program that monitors a region around the Milky Way's supermassive black hole, Sagittarius A* (Sgr A*). Four bright, variable X-ray sources were discovered within 3 lightyears of Sgr A*. Strong variability, which is present in all the sources, is indicative of an X-ray binary system where a black hole or neutron star is pulling matter from a nearby companion star. These relatively small, stellar-mass black holes, along with neutron stars, appear to have migrated into the Galactic center over the course of several billion years. Such a dense stellar graveyard has been predicted for years, and this represents the best evidence to date of its existence. Among the thousands of X-ray sources detected within 70 lightyears of Sgr A*, Munoz et al. [306] searched for those most likely to be active black holes and neutron stars by selecting only the brightest sources that also exhibited large variations in their X-ray output.

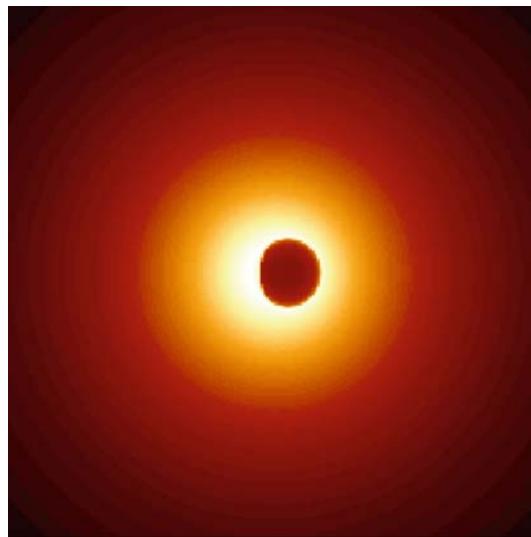


Fig. 9.10. Asymmetric cyclotron emission from an optically thin inner disk around Sgr A*. Depending on the inclination angle between the disk and the line-of-sight, only emission from the approaching side of the disk is boosted by the Doppler effect. To a distant observer, the event horizon casts a relatively large shadow with an apparent diameter of $\simeq 10$ gravitational radii that is due to the bending of light by the black hole, and this shadow is nearly independent of the black hole spin or orientation [151]. The predicted size ($\simeq 30 \mu\text{as}$) of this shadow for Sgr A* approaches the resolution of future radio interferometers. If the black hole is maximally spinning and viewed edge-on, then the shadow will be offset by $\simeq 8 \mu\text{as}$ from the center-of-mass and will be slightly flattened on one side. Image provided by A. Müller (LSW Heidelberg)

BH-Bulge Correlations in Nearby Galaxies

There is now good evidence for supermassive black holes in at least 50 nearby galaxies (Table 9.2). This is enough for a quantum improvement in what one can learn about black holes and galaxy formation. First, the amount of mass, that is found in black holes, is consistent with predictions of the waste mass left behind by quasars. Also, the individual masses of the black holes are consistent with predictions from quasar energies.

Two new results are fundamental correlations between black hole masses and the properties of their host galaxies. Galaxies come in two basic types, flat spinning disks and more nearly spherical bulges that rotate a little but that mostly are supported by random motions of stars. Many galaxies, like our own and the Andromeda Galaxy, consist of a bulge in the middle of a disk. When a galaxy contains only a bulge and not a disk, it is called an elliptical galaxy. In the following discussion, use of the term “bulge” includes elliptical galaxies. Supermassive black holes have now been found in elliptical galaxies and in galaxies that contain both a bulge and a disk, but not in galaxies that consist only of a disk. In 1993, John Kormendy found that black hole mass is roughly proportional to the luminosity of the bulge component of the host galaxy, which is a measure for the mass of the bulge. This is confirmed by the new black hole detections (Fig. 9.11). It implies that the mass of a black hole is always about 0.2% of the mass of the bulge

$$\log(M_H/M_\odot) = (8.20 \pm 0.1) + (1.12 \pm 0.06) \log(M_{\text{Bulge}}/10^{11} M_\odot). \quad (9.3)$$

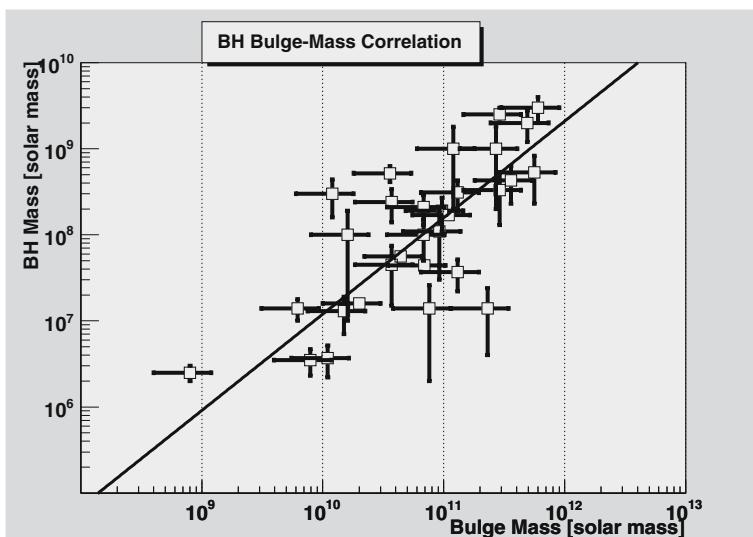


Fig. 9.11. Correlation between the bulge mass and black hole mass. For a total of 30 galaxies bulge masses were derived through Jeans equation modelling or adopted from dynamical models in the literature. The *solid line* gives the linear regression fit to the data with a slope of 1.12 ± 0.06 . This figure is based on data given in Table 9.2

The cause of this correlation is not known, but it implies that, as galaxies form, an approximately standard fraction of the mass ends up in the black hole. The correlation contains important clues to the origin and growth of galaxies.

In the last few years, a new and more fundamental correlation has been found. More massive black holes live in galaxies whose stars move faster (Fig. 9.12). This is often called the *Magorrian relation*, based on a sample of 26 galaxies, including 13 galaxies with new determinations of black hole masses from Hubble Space Telescope measurements of stellar kinematics [167]

$$\log(M_H/M_\odot) = \alpha + \beta \log(\sigma_e/200 \text{ km s}^{-1}). \quad (9.4)$$

The two parameters based on the data of Table 9.2 have the values $\alpha = 8.06 \pm 0.13$ and $\beta = 3.67 \pm 0.70$, while the original data lead to a steeper correlation with $\beta = 4.2$. Of course, the stars near the center must have high velocities; they are the ones that are used to find the black holes. Concluding that black holes correlate with these stars would be circular reasoning. Instead, the new correlation involves the stars in the main bodies of the galaxies. These stars do not feel the black holes. But they, too, move more rapidly than do stars in galaxies with less massive black holes. The scatter in the new correlation is almost zero. That is, it is almost the same as the measurement errors. Tight correlations in astronomy have always led to fundamental advances in our understanding of how things work. They tell us that there is an underlying astrophysical constraint that we did not know about before. In the present case, we do not yet have an explanation of why the correlation is so tight.

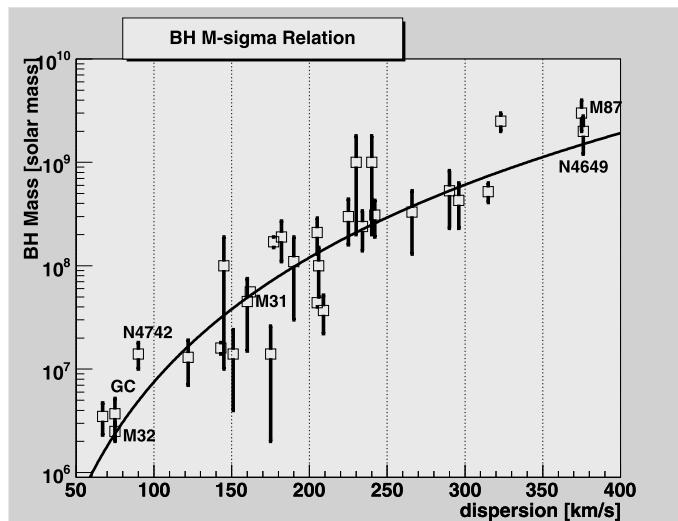


Fig. 9.12. Correlation between the bulge stellar dispersion and black hole mass, based on the data given in Table 9.2. The *solid line* corresponds to a scaling $M_H \propto \sigma_e^4$, normalized to the mass of M32. Note that the bulge of the Galactic center (GC) and the compact elliptical M32, a companion of M31, have identical bulge properties

But it implies that there is something almost magically regular about the process by which black holes are fed and grown.

In Fig. 9.13 we extend this diagram towards smaller masses to include objects such as globular clusters and dwarf ellipticals. If the Magorrian relation were to apply also for these objects, one would expect black hole masses in globular clusters of the order of at least a few hundred solar masses. Do dwarf elliptical and dwarf spiral galaxies contain central black holes with masses below 10^6 solar masses? Beyond the Local Group, dynamical searches for black holes in this mass range are very difficult, but the detection of accretion-powered nuclear activity could be used to infer the presence of a black hole. The nearby dwarf spiral galaxy NGC 4395 hosts a faint Seyfert 1 nucleus with a likely black hole mass in the range 10^4 – 10^5 solar masses, and for more than a decade it has been the only known example of a Seyfert 1 nucleus in a dwarf galaxy. Its velocity dispersion is 37 km/s, suggesting a possible black hole mass of order 10^5 solar masses. Another example is POX 52. While POX 52 was originally thought to be a Seyfert 2 galaxy, new data reveal an emission-line spectrum very similar to that of the dwarf Seyfert 1 galaxy NGC 4395, with clear broad components to the permitted line profiles. The host galaxy appears to be a dwarf elliptical. Applying scaling relations to estimate the black hole mass from the broad H_β linewidth and continuum luminosity, Barth et al. [60] find $M_H \simeq 1.6 \times 10^5 M_\odot$. The stellar velocity dispersion in the host galaxy is 36

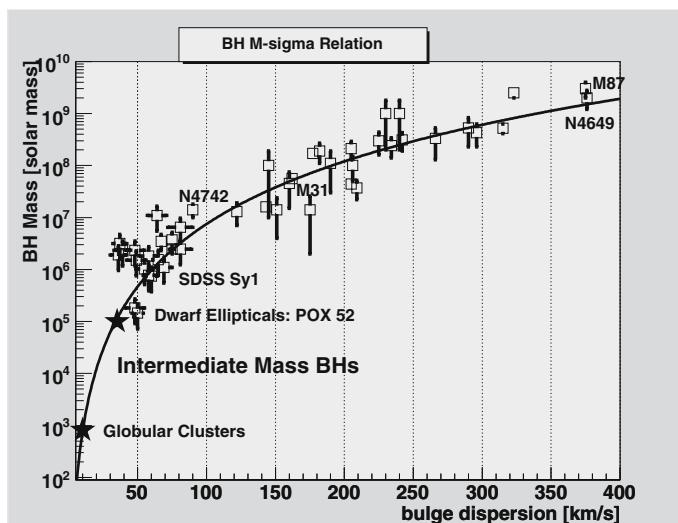


Fig. 9.13. Correlation between the bulge stellar dispersion and black hole mass, based on the data given in Table 9.2, extended towards smaller masses. The *solid line* corresponds to a scaling $M_H \propto \sigma_e^4$, normalized to the mass of M32. Based on this scaling, globular clusters should have black hole masses in the range of at least a few hundred solar masses. SDSS Sy1 denote Seyferts, which were selected from the Sloan Digital Sky Survey to have estimated black hole masses around 10^6 solar masses [60]

km/s, also suggestive of a black hole mass of order $10^5 M_\odot$. On the other hand, in M33 and NGC 205 only upper limits of $4 \times 10^4 M_\odot$ to black hole masses have been found.

9.2.4 Black Holes in Quasars

Since the pioneering works by Salpeter (1964), Lynden-Bell (1969) and Blandford and Rees (1974), the standard model for the nuclear activity in galaxies has envisaged a supermassive compact object, usually a black hole with mass 10^6 – $10^{10} M_\odot$, which acts as a central engine in the production of radio emission. In this picture, the radio power is derived from twin collimated beams of relativistic material which are ejected along the black hole spin axis. The total energy output is related to both the mass M_H of the central object and the accretion rate \dot{M} . The maximum source luminosity is set by the Eddington limit,

$$L_{\text{Ed}} = 1.3 \times 10^{47} \text{ erg s}^{-1} \frac{M_H}{10^9 M_\odot} \quad (9.5)$$

and the required accretion rate

$$\dot{M}_{\text{Ed}} \simeq 20 M_\odot \text{ yr}^{-1} \frac{0.1}{\epsilon_H} \frac{L}{10^{47} \text{ erg s}^{-1}}. \quad (9.6)$$

ϵ_H is the efficiency of mass–energy conversion. To produce powerful 3C radio galaxies and quasars with luminosities $L \simeq 10^{46}$ – 10^{47} erg s $^{-1}$ requires minimum black hole masses of $10^9 M_\odot$ and accretion rates of a few $M_\odot \text{ yr}^{-1}$ of infalling gas. The total amount of gas accreted in the lifetime of the quasar is similar to the black hole mass itself. This amount of gas is indeed available on the parsec scale of the host galaxies. The total amount of gas assembled in the core of these giant elliptical galaxies can be an appreciable fraction of the core mass itself.

The mass of a dead quasar can also be estimated from the total dissipation

$$M_H = \frac{L_Q \tau}{\epsilon_H c^2} = 7 \times 10^8 M_\odot \frac{L_Q}{10^{12} L_\odot} \frac{\tau}{10^9 \text{ yr}} \frac{0.1}{\epsilon_H}, \quad (9.7)$$

where L_Q is the quasar luminosity and τ its lifetime. 10^9 years is an upper limit to the lifetime. The density of such remnants follows from the integrated comoving energy density in quasar light

$$u \simeq 1.3 \times 10^{-15} \text{ erg cm}^{-3}. \quad (9.8)$$

The corresponding mass density is then a question of efficiency

$$\varrho_H = \frac{u}{\epsilon_H c^2} = 2.2 \times 10^5 M_\odot \text{ Mpc}^{-3} \frac{0.1}{\epsilon_H}. \quad (9.9)$$

On the other hand, the luminosity density of galaxies is $j \simeq 1.5 \times 10^8 L_\odot \text{ Mpc}^{-3}$, and a typical bright hot component contributes $8.5 \times 10^9 L_\odot$, depending somewhat

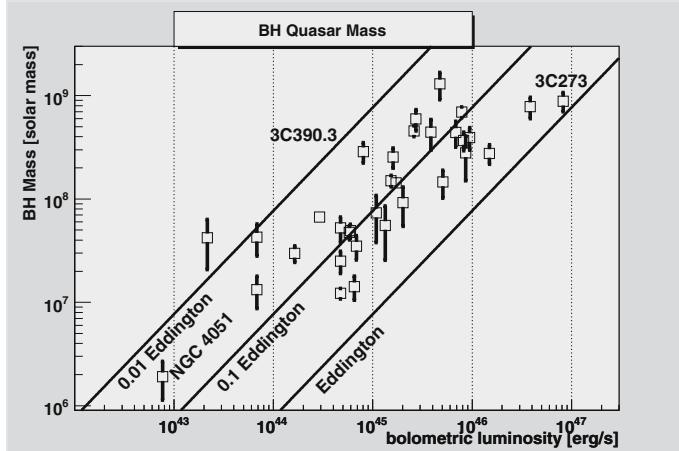


Fig. 9.14. Black hole masses for AGNs vs. luminosity as obtained by reverberation mapping. The luminosity scale on the lower axis is the bolometric luminosity $L_{\text{bol}} \simeq 9\lambda L_{\lambda}$. The diagonal lines show the Eddington limit L_{Edd} , $0.1 L_{\text{Edd}}$, and $0.01 L_{\text{Edd}}$. This figure is based on data from [328]

on the Hubble constant. This provides us with a number density of 2×10^{-2} hot galaxies Mpc^{-3} . The mean black hole mass required per hot galaxy is only $10^7 M_{\odot}$, which is in accordance with the observations.

If the emitting gas in the broad-line region of a quasar is assumed to be gravitationally confined, then reverberation mapping permits a simple dynamical determination of central mass. Correlated variability between the continuum and emission lines determines the radius of the broad-line region and the width of the line provides a measure of its velocity dispersion. This technique has been successfully used to determine BLR sizes for 17 Seyfert I galaxies, but quasar samples have been more difficult to monitor.

In Fig. 9.14, we present improved black hole masses for 35 active galactic nuclei (AGNs) based on a complete and consistent reanalysis of broad emission-line reverberation-mapping data [328]. From objects with multiple line measurements, one finds that the highest precision measure of the virial product is obtained by using the cross-correlation function centroid (as opposed to the cross-correlation function peak) for the time delay and the line dispersion (as opposed to full width half maximum) for the linewidth and by measuring the linewidth in the variable part of the spectrum. The scatter about the mass–luminosity relationship for these AGNs appears to be real and could be correlated with the Eddington ratio and object inclination. Typically, AGNs are not detectable as broadline objects for luminosities less than about 0.01 the Eddington luminosity.

9.3 Estimating Black Hole Spin

According to the two-hair theorem (Fig. 1.9), astrophysical black holes are given by two independent parameters: the mass M_H and the angular momentum J_H . As with neutron stars, we expect that accreting black holes are fast spinning objects. Rotating objects have in general an additional source of energy. The rotational energy of neutron stars is behind the powerful pair winds which are injected into surrounding nebulae (e.g. the Crab Nebula, Fig. 1.8). As we have seen, for black holes, energies add quadratically

$$M_H = \sqrt{M_{\text{irr}}^2 + \left(\frac{J_H}{2GM_{\text{irr}}/c} \right)^2}. \quad (9.10)$$

The irreducible mass M_{irr} is a kind of rest mass for a rotating black hole that cannot be further reduced by any physical processes, except Hawking radiation – which is however unimportant for macroscopic black holes. The second term is due to the rotational energy of black holes and is given in terms of the angular momentum J_H of the black hole. Therefore, each black hole contains a rotational energy

$$E_{\text{rot}} = (M_H - M_{\text{irr}})c^2 = M_H c^2 \left(1 - \sqrt{\frac{1}{2}(1 + \sqrt{1 - a_*^2})} \right), \quad (9.11)$$

which could be extracted by means of some electrodynamic processes. a_H denotes the specific angular momentum (Kerr parameter). Since $E_{\text{rot}} \leq 0.29 M_H c^2$, the rotational energy stored in the black hole of a quasar is a considerable amount of energy

$$E_{\text{rot}} < 5 \times 10^{55} \text{ watt s} \frac{M_H}{10^9 M_\odot}. \quad (9.12)$$

No other object can approach this upper limit for the rotational energy. The mass M_H of black holes residing in giant elliptical galaxies can easily exceed $10^9 M_\odot$. If this rotational energy could be dissipated in a kind of pulsar process, this would represent a considerable luminosity

$$L_{\text{rot}} \simeq \frac{E_{\text{rot}}}{t_{\text{diss}}} \simeq 1.6 \times 10^{40} \text{ watt} \frac{M_H}{10^9 M_\odot} \frac{10^9 \text{ yr}}{t_{\text{diss}}}, \quad (9.13)$$

essentially comparable to the mean luminosity of radio quasars at redshift of 2 – provided the system is able to dissipate the energy on a time-scale of a few billion years. This process could be behind the energetization of bright quasars, such as 3C 273 and 3C 345. In fact, Rees et al. had already suggested in 1982 that the nonthermal power of radio-loud objects (quasars and radio galaxies) could be accounted for by this rotational energy.

9.3.1 Black Hole Spin and Radio Galaxies

This energy-loss rate is an interesting number which can be compared to the bulk kinetic power Q_j in jets of radio galaxies and radio-loud quasars, estimated from the by-products of the jets, the large-scale radio lobes

$$Q_j = \frac{k U}{\tau_j}. \quad (9.14)$$

Here U is the energy stored in the lobes, taken from equipartition energy U_{eq} , $k \simeq 2$ allows for $P dV$ work expended by the jet on pushing back and warming up the extended medium. τ_j is the age of the jet, estimated, e.g. from the spectral age of the radio lobes. This amounts to a maximal jet power $\simeq 10^{40}$ watt for bright radio quasars and narrow-line radio galaxies for $0.5 < z < 1.0$. On the other hand, material radiatively excited by an AGN cools by line emission. Radio galaxies are usually only narrow-line emitters and one can therefore easily estimate the total narrow-line luminosity L_{NLR} in all narrow lines. Rawlings and Saunders (1991) obtained for an unbiased sample of FR II radio galaxies and low-power FR IIs a correlation between the jet power and L_{NLR} :

- $Q_j \propto L_{\text{NLR}}$ for FR I and FR II radio galaxies, as well as for radio quasars;
- $Q_j \simeq 100 L_{\text{NLR}} \simeq 10^{36}\text{--}10^{40}$ watt;
- radio-quiet quasars do not satisfy the $Q_j - L_{\text{NLR}}$ correlation.

The last point strongly indicates that the jet power is some extra power provided, e.g. by the rotational energy of the central source. But also radio-loud objects do have photoionizers which by virtue of $Q_j \propto L_{\text{NLR}}$ are controlled by the jet driving mechanism. It is also interesting that jet sources with given Q_j have a higher low-frequency luminosity when the sources are in a dense cluster environment. This probably indicates that the narrow line emission is due to the interaction of the jets with the ambient medium. Since the above correlation extends over more than four orders of magnitude, this could reflect the scaling of the central mass from $\simeq 10^6 M_\odot$ in faint ellipticals to $\simeq 10^{10} M_\odot$ in giant ellipticals. If jet power were related to rotational energy of the black hole, then its essentially only the mass and the angular momentum that dictate the jet power.

9.3.2 Spectral Fitting of Accretion Disks

When a black hole has a large accretion rate exceeding about a few percent of the Eddington rate, the accreting gas tends to be optically thick and to radiate approximately as a black-body (see Sect. 10.4). If the disk radiates as a true black-body at each radius, the total spectrum can easily be calculated and compared to the spectral fluxes observed at Earth. This flux will depend on $R_{\text{in}}^2 \cos i / d^2$, where i is the inclination angle, d the distance and R_{in} the inner edge of the disk. If the inner edge were at the ISCO, then one could get an estimate for the Kerr parameter a . These sources have to be in the high state. There is a major drawback in this method,

since a number of effects (opacity, Thomson scattering, etc.) will cause the spectrum of an accretion disk to deviate from a black-body. Therefore, spin estimates based on this method should be treated with caution.

9.3.3 Relativistic Iron Lines

Tanaka et al. [386] discovered for the first time a strong broad spectral line in the X-ray spectrum of the Seyfert galaxy MCG–6–30–15. They interpreted the line as fluorescent iron K α emission from cool gas in the accretion disk. Similar broad lines have been seen in various AGN and XRBs [341]. The rest energy of this iron line is at 6.4 keV, but the observed line profile extends from 4 to 7 keV (Fig. 1.1), which is the result from Doppler broadening and from gravitational redshifts (see Sect. 8.6.4). The line profile depends on several factors, such as the inner radius of the emitting region, the disk inclination, the emissivity of the disk material, and the motion of the emitting gas [303]. As the BH spin increases, the inner edge of the disk moves closer to the horizon and the velocity of the gas increases. As a result, the line extends down to very low energies (below 4 keV). This large redshift could be used as evidence for a high spin of the black hole. In the case of MCG–6–30–15, the line profile turned out to be highly time-variable. This is probably a sign that the inner edge of the standard disk is wobbling around with time, and the X-ray emission is highly turbulent due to the high level of turbulence in the inner hot disk. Among the XRBs, the source GX 339–4 shows a broad iron line with indication for a high spin, $a > 0.8$.

9.3.4 Quasiperiodic Oscillations

The power spectrum of intensity variations shows one or two peaks at frequencies around a few hundred Hz in the case of several BH XRBs [283] (e.g. as in Cyg X-1). The observed high frequency suggests that these oscillations arise in gas that is close to the horizon. This gas is probably strongly influenced by relativistic effects. However, there is at present no clear understanding for the origin of these oscillations.

9.4 Black Holes and Galaxy Formation

So far, astronomers have found a supermassive black hole in every galaxy observed that contains a bulge component. Therefore, the observed correlations say that black hole mass is intimately connected with bulge formation. Alternative theories come in two extremes:

- Black holes came first in a standard size, namely 0.2% of the mass of the first galaxy fragments. Then mergers of small galaxies made big galaxies, and the big galaxies still contained 0.2% mass black holes because, when two galaxies merge, their black holes merge, too.

- Black holes started out small and then grew during galaxy formation. If 0.2% of the gas that makes stars always gets fed to the central black hole, then the black hole mass fraction is always 0.2%.

Both theories include an explanation of quasars, but they differ in how they use quasars. In theory (i), the black holes come first and then regulate galaxy formation, while in theory (ii) black holes and galaxies grow together. Two arguments are in favor of the second scenario.

Observations show that there are two kinds of bulge-like components and that both contain black holes. One kind, called a pseudobulge, is believed to form in a bulgeless pure disk galaxy when gas flows inward toward the center. One observes that disks do not contain supermassive black holes in nearly the same proportion (0.2% of the mass) as do bulges. But seven galaxies contain pseudobulges and all of them contain standard black holes with about 0.2% of the pseudobulge mass. So the black holes must have grown during the process that made the pseudobulges.

The second argument comes from a comparison of the two correlations of black hole mass with host galaxy properties. The first correlation says that bigger bulges contain bigger black holes, with exceptions: a few galaxies contain anomalously big black holes. But the stars in these galaxies move faster, too, and they do so by precisely the right amount so that the scatter in the black hole mass – random velocity correlation is small. What does this mean? The reason why the stars move so rapidly is that the galaxy collapsed to an unusually small size when it formed. Then stars are closer together, so their gravitational forces on each other are bigger, so they must move faster. But, if black holes are unusually massive, whenever galaxies are unusually collapsed, then black hole masses were probably fixed by the collapse process. The alternative – that bigger black holes cause a galaxy to be more collapsed – is less likely, because bigger black holes would power brighter quasars; their radiation would push on the protogalactic gas and would tend to make it collapse less, not more.

Based on these arguments, one can conclude that the major events that made the bulge and the major periods of black hole growth were the same events. Galaxy formation directly results in the black hole feeding that makes quasars shine.

9.5 Black Hole Magnetospheres

Magnetic fields are essential ingredients in the physics of accretion disks around rapidly rotating black holes. Seed fields advected inwards from the ambient medium will be sheared into helical fields. Once these fields are advected into the neighborhood of the horizon, additional forces appear which are related to the spin of the black hole. These effects could amplify the fields and produce a dynamo action in a boundary layer between the horizon and the disk (Khanna and Camenzind [222]). In this way, the black hole could be immersed into a rotating magnetosphere.

The physics of black hole magnetospheres goes back to the pioneering work of Blandford and Znajek in 1977 [80]. These authors investigated the interaction of

a horizon with an external force-free magnetosphere and came to the conclusion that by means of this interaction rotational energy could be extracted out of the immense energy reservoir of a supermassive black hole. In the last 20 years, many authors worked on this problem which could be important for the launch of energetic jets in quasars, radio galaxies and microquasars. Most of these investigations were based on the force-free approximation, where plasma inertia is neglected in the Lorentz force. This force-free approximation is, however, not suitable for the discussion of plasma processes that are relevant in accretion and outflows. For this reason, we will present an introduction into the complete magnetohydrodynamics of rotating black holes including currents and plasma motion. This theory is quite complicated, when working however in the 3+1 split of Kerr spacetime the equations can be formulated in a way familiar to astrophysicists.

9.5.1 The 3+1 Formalism for Maxwell's Equations

The 3+1 formalism is of special importance for the discussion of electromagnetic fields in general relativity. This technique is based on a slicing of the spacetime such that the slices are space-like. This defines a vector field ∂_t which can be decomposed into normal and parallel components relative to the slicing

$$\partial_t = \alpha \mathbf{n} + \beta . \quad (9.15)$$

Here \mathbf{n} is the unit normal field and β is tangent to the slices. α is the lapse function (or redshift factor) and β the shift vector field. The line element of the Kerr geometry is a special decomposition of the general expression

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij} (dx^i + \beta^i dt) (dx^j + \beta^j dt) . \quad (9.16)$$

The four one-forms $\Theta^0 = \alpha dt$ and $\Theta^i = \theta^i + \beta^i dt$ ($i = 1, 2, 3$) are therefore orthonormal and form a natural basis for one-forms. θ^i are normalized one-forms in the spatial slice. In Kerr space, α is explicitly known, and due to axisymmetry the shift vector has only one nonvanishing component $\beta^\phi = -\omega$. In the case of rapidly rotating neutron stars, these two functions are only given numerically.

The derivation of Maxwell's equations adapted to the slicing is quite complicated when usual techniques are involved. There is a very elegant derivation based on the technique of differential forms, the electromagnetic potential $A = A_\mu dx^\mu$ is essentially a one-form and the Faraday tensor F a two-form. Instead of using the coordinate basis, the Faraday tensor is now decomposed with respect to this natural basis of one-forms

$$F = E \wedge \Theta^0 + B \quad (9.17)$$

with the definition of the two fields

$$E = E_i \Theta^i , \quad B = \frac{1}{2} B_{ij} \Theta^i \wedge \Theta^j . \quad (9.18)$$

On the other hand, it is useful to decompose these quantities also into components adapted to the slicing

$$E = \mathcal{E} + (E_i \theta^i) dt = \mathcal{E} + i_\beta \mathcal{E} dt. \quad (9.19)$$

$\mathcal{E} = E_i \theta^i$ is now the *horizontal electric field*. Similarly, we have for the magnetic two-form

$$\mathcal{B} = \frac{1}{2} B_{ij} \theta^i \wedge \theta^j. \quad (9.20)$$

We obtain therefore the following decomposition of the Faraday two-form

$$F = \mathcal{B} + (\alpha \mathcal{E} - i_\beta \mathcal{B}) \wedge dt. \quad (9.21)$$

Since the homogeneous Maxwell's equations are given as $dF = 0$, this means

$$d\mathcal{B} + \partial_t \mathcal{B} \wedge dt + d(\alpha \mathcal{E}) \wedge dt - d(i_\beta \mathcal{B}) \wedge dt = 0. \quad (9.22)$$

This splits into the two equations

$$d\mathcal{B} = 0, \quad d(\alpha \mathcal{E}) + \partial_t \mathcal{B} = d(i_\beta \mathcal{B}). \quad (9.23)$$

Using the Cartan formula for the Lie derivative, $\mathcal{L}_\beta = di_\beta + i_\beta d$, the homogeneous Maxwell's equations can also be written as

$$d\mathcal{B} = 0, \quad d(\alpha \mathcal{E}) + \partial_t \mathcal{B} = \mathcal{L}_\beta \mathcal{B}. \quad (9.24)$$

The inhomogeneous equations follow from the Hodge dual one-forms

$$\mathcal{H} = * \mathcal{B}, \quad \mathcal{E} = * \mathcal{D}. \quad (9.25)$$

(* denotes the three-dimensional Hodge dual). The dual of the Faraday two-form is therefore

$$*F = \mathcal{D} - (\alpha \mathcal{H} + i_\beta \mathcal{D}) \wedge dt. \quad (9.26)$$

In the same sense we can also decompose the current density

$$J = \varrho_e \theta^0 + j_k \theta^k, \quad (9.27)$$

where ϱ_e is the electric charge density with respect to this basis. We also introduce the horizontal forms

$$\mathbf{q} = \varrho_e \theta^1 \wedge \theta^2 \wedge \theta^3, \quad \mathbf{j} = j_k \theta^k, \quad \mathcal{J} = * \mathbf{j}. \quad (9.28)$$

From this we find for the current three-form

$$*J = \mathbf{q} + (i_\beta \mathbf{q} - \alpha \mathcal{J}) \wedge dt. \quad (9.29)$$

With this, the inhomogeneous Maxwell's equations can be written

$$\begin{aligned} d * F &= d\mathcal{D} + \partial_t \mathcal{D} \wedge dt - d(\alpha \mathcal{H}) \wedge dt - d(i_\beta \mathcal{D}) \wedge dt \\ &= 4\pi \mathbf{q} + 4\pi(i_\beta \mathbf{q} - \alpha \mathbf{J}) \wedge dt. \end{aligned} \quad (9.30)$$

This once again splits into two equations

$$d\mathcal{D} = 4\pi \mathbf{q}, \quad d(\alpha \mathcal{H}) = (\partial_t - \mathcal{L}_\beta) \mathcal{D} + 4\pi \alpha \mathbf{J}. \quad (9.31)$$

The first equation is Gauss' law, and the second one Ampère's law modified by a Lie derivative along the shift vector field. These are Maxwell's equations adapted to the slicing of any geometry.

This equation also implies charge conservation in the form of

$$(\partial_t - \mathcal{L}_\beta) \mathbf{q} + d(\alpha \mathbf{J}) = 0. \quad (9.32)$$

The Hodge dual of this form is equivalent to the usual form of charge conservation

$$(\partial_t - \mathcal{L}_\beta) \varrho_e + \nabla \cdot (\alpha * \mathbf{J}) - \alpha \varrho_e \text{Tr}(\mathbf{K}) = 0. \quad (9.33)$$

\mathbf{K} denotes the second fundamental form of the slices defined as

$$K_{ij} = \frac{1}{2\alpha} (\beta_{i|j} + \beta_{j|i}). \quad (9.34)$$

Due to axisymmetry, we have in Kerr space $\text{Tr}(\mathbf{K}) = 0$. This corresponds in fact to the charge conservation law discussed in Thorne et al. [392]

$$\alpha \frac{d\varrho_e}{d\tau} = (\partial_t - \mathbf{B} \cdot \nabla) \varrho_e = -\nabla \cdot (\alpha \mathbf{j}). \quad (9.35)$$

Maxwell's Equations under Axisymmetry

In the following, we will not work with forms, but with the corresponding vector fields. \mathbf{E} and \mathbf{B} denote the electric and magnetic fields, respectively, as measured by ZAMOs, which are special observers having a four-velocity U perpendicular to the absolute space

$$U = \frac{1}{\alpha} (\partial_t - \beta^i \mathbf{e}_i). \quad (9.36)$$

Together with the orthonormal tetrad in Boyer–Lindquist coordinates \mathbf{e}_r , \mathbf{e}_θ and $\mathbf{e}_\phi = (1/\tilde{\omega}) \partial_\phi$ this forms a physical basis in the tangential space. This is the basis that is dual to the above natural one-form basis. The circumference of a circle around the rotational axis is measured by $\tilde{\omega} = \sqrt{g_{\phi\phi}}$. The four-velocity of the plasma u is then expressed as

$$u = \gamma(U + \mathbf{v}), \quad (9.37)$$

where \mathbf{v} is now the three-velocity of the plasma with respect to ZAMOs.

Instead of using the forms, Maxwell's equations (9.24) and (9.31) are expressed for the vector fields \mathbf{E} and \mathbf{B} , which live on the slice,

$$\nabla \cdot \mathbf{E} = 4\pi\varrho_e, \quad \nabla \cdot \mathbf{B} = 0 \quad (9.38)$$

$$\nabla \times (\alpha \mathbf{E}) = -(\partial_t - \mathcal{L}_\beta) \mathbf{B} \quad (9.39)$$

$$\nabla \times (\alpha \mathbf{B}) = (\partial_t - \mathcal{L}_\beta) \mathbf{E} + 4\pi\alpha \mathbf{j}. \quad (9.40)$$

The current density \mathbf{j} is given by Ohm's law,

$$\mathbf{j} = \sigma\gamma (\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \varrho'_e \gamma \mathbf{v}, \quad (9.41)$$

where σ denotes the conductivity, \mathbf{v} the bulk velocity of the plasma, γ the corresponding Lorentz factor, and ϱ'_e the charge density in the rest frame of the plasma. As compared to flat spacetimes, there are two important additional terms related to the shift vector of the slicing (or the frame-dragging effect in Kerr space).

For axisymmetric fields, Maxwell's equations assume the simple form

$$\nabla \times (\alpha \mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{B} \cdot \nabla \omega) \tilde{\omega} \mathbf{e}_\phi \quad (9.42)$$

$$\nabla \times (\alpha \mathbf{B}) = \frac{\partial \mathbf{E}}{\partial t} - (\mathbf{E} \cdot \nabla \omega) \tilde{\omega} \mathbf{e}_\phi + 4\pi\alpha \mathbf{j}. \quad (9.43)$$

On an axisymmetric spacetime, it is now useful to split all the vector fields into poloidal and toroidal components, $\mathbf{B} = \mathbf{B}_p + \mathbf{B}_T$ with $\mathbf{B}_T = B^\phi \mathbf{e}_\phi$ and $\mathbf{B}_p = B^r \mathbf{e}_r + B^\theta \mathbf{e}_\theta$, and $\mathbf{B}_p = \nabla \times (A^\phi \mathbf{e}_\phi)$. The induction and Ampère's equations give the following relations, when decomposed into poloidal and toroidal components

$$\frac{\partial A^\phi}{\partial t} = -\alpha E^\phi \quad (9.44)$$

$$\frac{\partial B^\phi}{\partial t} = \tilde{\omega} \mathbf{B}_p \cdot \nabla \omega - \mathbf{e}_\phi \cdot (\nabla \times \alpha \mathbf{E}_p) \quad (9.45)$$

$$\frac{\partial E^\phi}{\partial t} = \tilde{\omega} \mathbf{E}_p \cdot \nabla \omega - \mathcal{G}_2[A^\phi] - 4\pi\alpha j^\phi \quad (9.46)$$

$$\frac{\partial \mathbf{E}_p}{\partial t} = \nabla \times (\alpha B^\phi) - 4\pi\alpha \mathbf{j}_p. \quad (9.47)$$

$\mathcal{G}_2[A] \equiv -\mathbf{e}_\phi \cdot [\nabla \times \alpha \mathbf{B}_p]$ is the Grad-Shafranov operator for the poloidal flux function $\Psi \equiv \tilde{\omega} A^\phi = A_\alpha m^\alpha$, explicitly given by divergence operator

$$\mathcal{G}_2[\Psi] \equiv \tilde{\omega} \mathbf{Div} \left[\frac{\alpha}{\tilde{\omega}^2} \nabla \Psi \right], \quad (9.48)$$

where **Div** is the divergence operator on three-space. In terms of this flux function Ψ , the poloidal magnetic field is given in the standard form

$$\mathbf{B}_p = \frac{1}{\tilde{\omega}} \nabla \Psi \times \mathbf{e}_\phi. \quad (9.49)$$

In addition, Ohm's law has the two components

$$\mathbf{j}_p = \sigma\gamma [\mathbf{E}_p + \mathbf{v}_T \times \mathbf{B}_p + \mathbf{v}_p \times \mathbf{B}_T] \quad (9.50)$$

$$j^{\hat{\phi}} = \sigma\gamma [E^{\hat{\phi}} + \mathbf{e}_{\phi} \cdot (\mathbf{v}_p \times \mathbf{B}_p)], \quad (9.51)$$

provided the charge density in the plasma frame vanishes. This formulation of Maxwell's equations is suitable for implementing in a code. Due to the potential formulation, $\text{Div}(\mathbf{B}) = 0$ is automatically satisfied. We will discuss in the following some consequences of the above equations.

The Event Horizon as a Membrane

Similar to neutron stars, boundary conditions must also be specified at the event horizon, at least in the vacuum case. In contrast to a neutron star, a black hole has no material surface. One can think of the black hole as having a fictitious surface charge density σ^H that compensates for the flux of electric field across the surface, and a fictitious surface current density \mathbf{j}^H that closes tangential components of the magnetic fields. This interpretation is the basis of the membrane formalism (Thorne et al. [392]).

Gauss' law implies for the electric field perpendicular to the horizon

$$E_{\perp}^H \rightarrow 4\pi\sigma^H, \quad (9.52)$$

and Ampère's law for the field parallel to the horizon

$$\alpha\mathbf{B}_{\parallel} \rightarrow \mathbf{B}_{\parallel}^H = 4\pi\mathbf{j}^H \times \mathbf{n}. \quad (9.53)$$

In addition, Ohm's law implies

$$\alpha\mathbf{E}_{\parallel} \rightarrow \mathbf{E}_{\parallel}^H = \mathcal{R}^H \mathbf{j}^H \quad (9.54)$$

with $\mathcal{R}^H = 4\pi/c = 377$ Ohm as the effective surface resistance of the horizon. The fields \mathbf{E}_{\perp} and \mathbf{B}_{\perp} are finite at the horizon, \mathbf{E}_{\parallel} and \mathbf{B}_{\parallel} generally diverge as $1/\alpha$, but

$$|\mathbf{E}_{\parallel} - \mathbf{n} \times \mathbf{B}_{\parallel}| \propto \alpha \rightarrow 0 \quad (9.55)$$

at the horizon. This signifies that for a locally nonrotating observer the electromagnetic field at the horizon looks like a wave sinking into the black hole (ingoing wave conditions). This formalism is now superseded by using Kerr–Schild coordinates.

9.5.2 Plasma Equations in the 3+1 Split

Besides the time evolution of magnetic fields we also need a description of the time evolution of the plasma part. We decompose the energy–momentum tensor

T into horizontal and vertical components (one-component description for the plasma)

$$T = \epsilon U \otimes U + U \otimes \overset{\leftrightarrow}{t} . \quad (9.56)$$

ϵ is the energy-density with respect to ZAMOs, \mathbf{S} the momentum flux and $\overset{\leftrightarrow}{t}$ the stress tensor. For a one-component nonviscous plasma the energy-momentum tensor has the form

$$T = (\varrho + P) u \otimes u + Pg + T_{\text{em}} \quad (9.57)$$

With the decomposition of $u = \gamma(U + \mathbf{v})$ we find the following components with respect to ZAMOs

$$\epsilon = \gamma^2 (\varrho + Pv^2) + \epsilon_{\text{em}} \quad (9.58)$$

$$\mathbf{S} = (\varrho + P) \gamma^2 \mathbf{v} + \mathbf{S}_{\text{em}} = \varrho_0 h \gamma^2 \mathbf{v} + \mathbf{S}_{\text{em}} \quad (9.59)$$

$$\overset{\leftrightarrow}{t} = (\varrho + P) \gamma^2 \mathbf{v} \otimes \mathbf{v} + P \overset{\leftrightarrow}{g} + \overset{\leftrightarrow}{t}_{\text{em}} = \mathbf{S} \otimes \mathbf{v} + P \overset{\leftrightarrow}{g} + \overset{\leftrightarrow}{t}_{\text{em}} . \quad (9.60)$$

ϱ_0 is the rest-mass density and $h = (\varrho + P)/n$ the relativistic specific enthalpy. \mathbf{S}_{em} represents the Poynting flux measured by ZAMOs and $\overset{\leftrightarrow}{t}_{\text{em}}$ the Maxwell stresses

$$\epsilon_{\text{em}} = \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2) \quad (9.61)$$

$$\mathbf{S}_{\text{em}} = \frac{1}{4\pi} \mathbf{E} \times \mathbf{B} \quad (9.62)$$

$$\overset{\leftrightarrow}{t}_{\text{em}} = \frac{1}{4\pi} \left(-\mathbf{E} \otimes \mathbf{E} - \mathbf{B} \otimes \mathbf{B} + \frac{1}{2} \overset{\leftrightarrow}{g} (\mathbf{E}^2 + \mathbf{B}^2) \right) . \quad (9.63)$$

Using the 3+1 split of the connection of Kerr space, one can now derive the 3+1 split of the hydrodynamic equations $\nabla \cdot T = 0$ (for more details, see Durrer and Straumann 1988 [139]). Energy conservation, given by $U \cdot (\nabla \cdot T) = 0$, can be written as

$$\frac{d\epsilon}{d\tau} = \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \epsilon = -\frac{1}{\alpha^2} \mathbf{Div}(\alpha^2 \mathbf{S}) + \epsilon \text{Tr} \left(\overset{\leftrightarrow}{K} \right) + \text{Tr} \left(\overset{\leftrightarrow}{K} \cdot \overset{\leftrightarrow}{t} \right) . \quad (9.64)$$

The first term in the energy conservation is the familiar divergence of the energy flux, with one factor of α inside the divergence to account for the gravitational redshift of the energy and the other to convert the per unit proper time in the definition of the flux into to a per unit universal time. The last term is interesting and can be written in terms of the shear $\sigma_{ik}^K \equiv -K_{ik}$ of adjacent observers

$$\frac{1}{\alpha} (\partial_t - \beta \cdot \nabla) \epsilon = -\frac{1}{\alpha^2} \mathbf{Div}(\alpha^2 \mathbf{S}) - \epsilon \text{Tr} \left(\overset{\leftrightarrow}{\sigma} \right) - \sigma_{ik}^K t^{ik} .$$

(9.65)

Similarly, Euler's equations, given by $\mathbf{h} \cdot (\nabla \cdot \mathbf{T}) = 0$, assume the form

$$\boxed{\frac{d\mathbf{S}}{d\tau} = \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \mathbf{S} = -\epsilon \nabla (\ln \alpha) - \frac{1}{\alpha} \mathbf{Div} \left(\alpha \overset{\leftrightarrow}{t} \right) + \text{Tr} \left(\overset{\leftrightarrow}{K} \right) \mathbf{S} + 2 \overset{\leftrightarrow}{K} \cdot \mathbf{S}.} \quad (9.66)$$

This can be brought into the form

$$\frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) \mathbf{S} = -\epsilon \nabla (\ln \alpha) - \frac{1}{\alpha} \mathbf{Div} \left(\alpha \overset{\leftrightarrow}{t} \right) + \text{Tr} \left(\overset{\leftrightarrow}{K} \right) \mathbf{S} - \boldsymbol{\sigma} \cdot \mathbf{S} - \frac{1}{2} \mathbf{H} \times \mathbf{S}. \quad (9.67)$$

$-\nabla \ln \alpha$ represents the local gravitational force measured by ZAMOs. For slicings with nonvanishing fundamental form, additional couplings occur between the curvature of the absolute space and the momentum flux. The term $-(1/2)\mathbf{H} \times \mathbf{S}$ is familiar from the precession law for gyroscopes; the momentum density \mathbf{S} will precess relative to absolute space.

For stationary flows on Kerr space this implies the two equations

$$\mathbf{Div}(\alpha^2 \mathbf{S}) = \alpha^2 \sigma_{ik}^K t^{ik} \quad (9.68)$$

$$\frac{1}{\alpha} \nabla_k (\alpha t_i^k) = -\epsilon \nabla_i (\ln \alpha) - \frac{1}{\alpha} S_\phi \nabla_i \omega. \quad (9.69)$$

In contrast to flat spaces, the energy flow is no longer conserved; the gravitomagnetic field, represented by the shear σ_{ik}^K , changes energy conservation. Even in the case of pure hydrodynamic disk accretion, this coupling between the plasma stress tensor, $t_{r\phi}$ (angular momentum flow vector), and the gravitomagnetic field introduces a work done on the disk plasma, which is very similar to viscous dissipation, $\sigma_{r\phi}^K t^{r\phi}$, except that here we have a coupling between the shear of absolute space and the momentum flux. This term can be very important if strong magnetic fields occur near the horizon of a rapidly rotating black hole.

9.5.3 Time Evolution of Magnetic and Current Flux in Turbulent Disks

The above formulation can now be used to investigate the time evolution of magnetic flux in accretion disks around rapidly rotating objects. This is a fundamental topic for the understanding of the formation of magnetospheres around rotating black holes.

The Grad–Shafranov Equation

When we combine Eq. (9.46) with the first equation (9.44) of Maxwell's equations, this provides us a kind of wave equation for the poloidal flux Ψ

$$\boxed{\frac{\partial^2 \Psi}{\partial t^2} - \alpha \tilde{\omega} \mathcal{G}_2[\Psi] = -\alpha \tilde{\omega}^2 \mathbf{E}_p \cdot \nabla \omega + 4\pi \tilde{\omega} j^\phi.} \quad (9.70)$$

This shows explicitly that the equation is hyperbolic, as required by Maxwell's theory. There are two source terms: (i) External toroidal currents provide a source for poloidal magnetic fluxes, as we already know from flat space electrodynamics. (ii) Differential rotation of absolute space, represented by $\nabla\omega$, shears poloidal electric fields which act then as an additional source for poloidal fluxes. This is especially interesting for *stationary configurations* where we find the *Grad-Shafranov equation* for the poloidal magnetic flux in axisymmetric space-times

$$\alpha \mathcal{G}_2[\Psi] = \alpha \tilde{\omega} \mathbf{E}_p \cdot \nabla \omega - 4\pi j^\phi. \quad (9.71)$$

This represents the general relativistic form of the flat Grad-Shafranov equation, $\alpha = 1$,

$$\mathbf{Div} \left(\frac{1}{\tilde{\omega}^2} \nabla \Psi \right) = -4\pi j^\phi. \quad (9.72)$$

Using Ohm's law, equation (9.70) can be rewritten as

$$\begin{aligned} \frac{\partial^2 \Psi}{\partial t^2} + 4\pi\gamma\sigma \alpha \frac{\partial \Psi}{\partial t} - \alpha \tilde{\omega} \mathcal{G}_2[\Psi] \\ = -\tilde{\omega}^2 \alpha \mathbf{E}_p \cdot \nabla \omega + 4\pi\alpha^2 \gamma\sigma \tilde{\omega} \mathbf{e}_\phi \cdot (\mathbf{v}_p \times \mathbf{B}_p). \end{aligned} \quad (9.73)$$

Using the expression for the poloidal magnetic field and the definition of the magnetic diffusivity $\eta = c^2/4\pi\sigma$, we obtain the equation

$$\begin{aligned} \eta \frac{\partial^2 \Psi}{c^2 \partial t^2} + \alpha\gamma \frac{\partial \Psi}{\partial t} + \alpha\gamma (\alpha \mathbf{v}_p \cdot \nabla) \Psi \\ - \eta \alpha \tilde{\omega} \mathcal{G}_2[\Psi] = -\eta \tilde{\omega}^2 \alpha \mathbf{E}_p \cdot \nabla \omega. \end{aligned} \quad (9.74)$$

The relevance of the first two terms in a time-dependent evolution of magnetic flux can be estimated by the ansatz $\Psi(t, r, \theta) = \exp(\Gamma t) \tilde{\Psi}(r, \theta)$. This yields the eigenvalue equation

$$\begin{aligned} \left(\frac{\Gamma^2}{4\pi\sigma} + \alpha\gamma\Gamma \right) \tilde{\Psi} + \alpha\gamma (\alpha \mathbf{v}_p \cdot \nabla) \tilde{\Psi} \\ - \eta \alpha \tilde{\omega} \mathcal{G}_2[\tilde{\Psi}] = -\eta \tilde{\omega}^2 \exp(-\Gamma t) \alpha \mathbf{E}_p \cdot \nabla \omega. \end{aligned} \quad (9.75)$$

The second derivative in time is in general $\Gamma/4\pi\sigma$ times the second term, which is first order in the time derivative. We can neglect second-order time derivatives, whenever the growth rates Γ are much less than the microscopic scales given by the conductivity $\sigma = 3.2 \text{ MHz } T_e^{3/2}$. In a turbulent plasma, the effective conductivity is, however, much less, and the corresponding growth times $1/\Gamma$ are of the order of the Alfvén transit time R/v_A at the radius R in a disk.

Causality requires that the second derivative is present in this transport equation for the poloidal magnetic flux. This term regulates the relaxation. For long term evolution we may neglect this part and end up with a parabolic diffusion type equation used in some simulations of the time evolution of magnetic fields in turbulent disks (Khanna and Camenzind [222])

$$\begin{aligned} \gamma \frac{\partial \Psi}{\partial t} + \gamma (\alpha \mathbf{v}_p \cdot \nabla) \Psi \\ - \eta \tilde{\omega} \mathcal{G}_2[\Psi] = -\eta \tilde{\omega}^2 \mathbf{E}_p \cdot \nabla \omega. \end{aligned} \quad (9.76)$$

Except for the term on the right-hand side, this is the standard equation for the evolution of magnetic flux in Newtonian magnetohydrodynamics.

The Current Flux Equation

The second equation determines the time evolution of the current function $T(t, r, \theta) = 2 \int \alpha \mathbf{j}_p \cdot d\mathbf{A} = \alpha \tilde{\omega} B^\phi(t, r, \theta)$. From equation (9.45) we get

$$\frac{\partial T}{\partial t} = \alpha \tilde{\omega}^2 \mathbf{B}_p \cdot \nabla \omega - \alpha \tilde{\omega} \mathbf{e}_\phi \cdot \left[\nabla \times \alpha \left(\frac{1}{\sigma \gamma} \mathbf{j}_p - \mathbf{v}_T \times \mathbf{B}_p - \mathbf{v}_p \times \mathbf{B}_T \right) \right], \quad (9.77)$$

when Ohm's law is used. Together with the poloidal component of Ampère's law this can be written as

$$\begin{aligned} \frac{\partial T}{\partial t} - \alpha \tilde{\omega} \mathbf{e}_\phi \cdot \left[\nabla \times \frac{\eta}{\gamma} (\nabla \times \alpha \mathbf{B}_T) \right] &= \alpha \tilde{\omega}^2 \mathbf{B}_p \cdot \nabla \omega \\ + \alpha \tilde{\omega} \mathbf{e}_\phi \cdot \nabla \times \left[\frac{\eta}{\gamma} \frac{\partial \mathbf{E}_p}{\partial t} - \mathbf{e}_\phi \cdot (\mathbf{v}_T \times \mathbf{B}_p - \mathbf{v}_p \times \mathbf{B}_T) \right]. \end{aligned} \quad (9.78)$$

This finally leads to a diffusion type equation for the current function

$$\begin{aligned} \frac{\partial T}{\partial t} + \alpha (\mathbf{v}_p \cdot \nabla) T - \alpha \tilde{\omega}^2 \nabla \cdot \left(\frac{T}{\tilde{\omega}^2} \mathbf{v}_p \right) - \alpha \tilde{\omega}^2 \nabla \cdot \left(\frac{\eta}{\gamma \tilde{\omega}^2} \nabla T \right) \\ = \alpha \tilde{\omega}^2 \mathbf{B}_p \cdot \nabla \Omega + \alpha \tilde{\omega} \mathbf{e}_\phi \cdot \nabla \times \left(\frac{\eta}{\gamma} \frac{\partial \mathbf{E}_p}{\partial t} \right). \end{aligned} \quad (9.79)$$

The left-hand side of this equation is nothing than the flat space expression, which tells us that toroidal magnetic fields change with time by means of advection, compression and diffusion. As in flat space, the main driving source for toroidal fields is the shearing of poloidal magnetic fields by the differential rotation of the plasma motion. The only new source term is due to induction by electric fields.

In contrast to equation (9.70), differential rotation of absolute space is only an indirect source for toroidal magnetic fields. Even for vanishing specific angular momentum of the plasma, $\lambda = 0$, an effective shear is produced by the differential

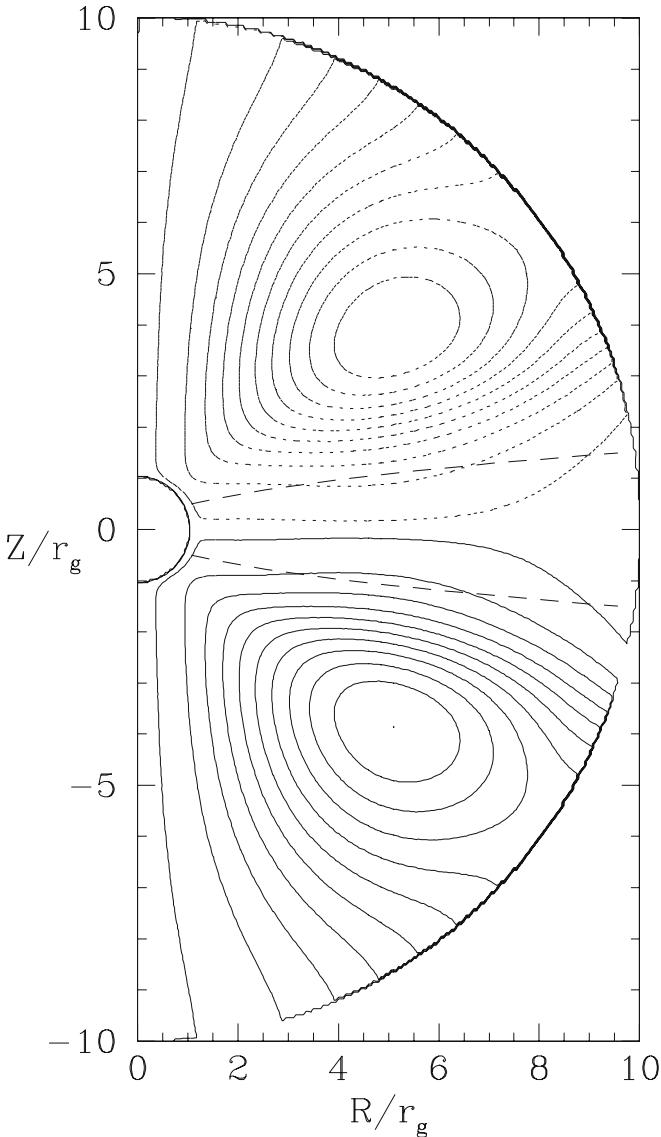


Fig. 9.15. Quadrupolar field topology for the toroidal field around a rapidly rotating black hole for $a_H = 0.998M_H$. This poloidal current flux distribution is generated by poloidal dipolar magnetic fields accreted onto a black hole. In this simulation, differential rotation is only driven by absolute space, i.e. the specific angular momentum of the plasma vanishes. The dashed lines indicate the scale-height of the disk. This topology leads to current loops which can power disk winds from the inner part of the accretion disk. In this way, rotational energy stored in the spacetime of a black hole is drained into Poynting flux which further out accelerates plasma to high Lorentz factors. Figure adapted from [99]

rotation of absolute space. A solution of this equation is shown in Fig. 9.15 for a dipolar poloidal field structure around a rapidly rotating black hole for a rotation law of the form (Camenzind and Khanna [99])

$$\Omega = \omega + \frac{\alpha^2}{R^2} \frac{\lambda}{1 - \lambda\omega}. \quad (9.80)$$

Since the hyperbolic nature of this equation is not obvious, it can be written in an alternative way by starting with the time derivative of the toroidal component B_T of the induction equation (9.46)

$$\frac{\partial^2 B_T}{\partial t^2} - \alpha \mathcal{J}_2[B_T] = \tilde{\omega} \frac{\partial \mathbf{B}_p}{\partial t} \cdot \nabla \omega + 4\pi \mathbf{e}_\phi \cdot [\nabla \times (\alpha^2 \mathbf{j}_p)]. \quad (9.81)$$

This shows explicitly the hyperbolic nature of the equation. Similar to the discussion of the transport of poloidal flux we may neglect the current displacement term in equation (9.79) which then yields a parabolic equation for the time evolution of the current function.

On the Validity of Cowling's Theorem in Kerr Geometry

Axisymmetric magnetic fields can generally be decomposed into a poloidal field and an azimuthal component. As we have seen, differential rotation can stretch the poloidal field and convert it into an azimuthal field. This gives an efficient amplification of the magnetic field, but it requires the poloidal field as a source. In strictly two-dimensional geometry, there is no corresponding source for the poloidal field, which must eventually decay. The flow can only advect the poloidal field around, but, because of axisymmetry, it cannot stretch or amplify the poloidal field. This is known as Cowling's antidynamo theorem.

Cowling's theorem therefore states that stationary axisymmetric magnetic fields cannot be maintained by purely axisymmetric plasma motions. In the Kerr geometry, near a rotating black hole, the situation is different. The hole's rotation drags all physical objects near it into orbital motion in the same direction as the hole rotates. Physical quantities, such as electric and magnetic fields, are measured by fiducial observers (ZAMOs), who are orbiting with angular velocity $\omega(r)$. This frame-dragging effect acts in a similar way as ordinary differential rotation in nonrelativistic hydromagnetics and leads to field-line stretching. The gravitomagnetic effect in the induction equation changes this conclusion of classical electrodynamics (Khanna and Camenzind [99, 222]). The reason is that in Kerr space the coupling between the gravitomagnetic shear $\nabla\omega$ and the poloidal electric field \mathbf{E}_p is a potential source for the poloidal magnetic flux Ψ (see equation (9.70)). This term is absent in nonrotating space and leads to the usual formulation of Cowling's theorem. In order to see this effect, one has to simulate the two equations, there are no analytical results known.

The kinematic evolution of axisymmetric magnetic fields in rotating magnetospheres of relativistic compact objects has also been analytically studied by Tomimatsu [395], based on relativistic Ohm's law in stationary axisymmetric geometry.

The first process is caused by the help of a background uniform magnetic field in addition to the dragging of inertial frames. It is shown that excited multipolar components of poloidal and azimuthal fields are sustained as stationary modes, and outgoing Poynting flux converges toward the rotation axis. The second process is a self-excited dynamo through azimuthal convection current, which is found to be effective if plasma rotation becomes highly relativistic with a sharp gradient in the angular velocity. In this case, no frame-dragging effect is needed, and the coupling between charge separation and plasma rotation becomes important.

Magnetic Engulfment of Black Holes

As discussed above, one can neglect for the long-term time evolution of magnetic fields in accretion disks the influence of the displacement current in the two equations (9.76) and (9.79). This leads then to a coupled parabolic system of equations for the poloidal magnetic flux $\Psi(t, r, \theta)$ and the poloidal current function $T(t, r, \theta)$ (Khanna and Camenzind [222]). The accretion profile with plasma rotation $\Omega(r)$ and accretion drift $v_r(r)$ must be given by some reasonable approximation, as following from the accretion disk solutions. Similarly, the magnetic diffusivity η is given by some characteristic turbulent diffusivity η_0 .

The two equations for Ψ and T are brought to dimensionless form by scaling radii in units of M_H , velocities in units of the speed of light, and time in units of the diffusion time $t_{\text{diff}} = M_H^2/\eta_0$. The advective terms are then scaled by a magnetic Reynolds number defined as

$$\mathcal{R}_m = \frac{M_H c}{\eta_0} = \frac{t_{\text{diff}} c}{M_H} \simeq \frac{M_H^2}{\alpha_T H^2} \gg 1, \quad (9.82)$$

where H is the disk scale-height at the horizon and $\alpha_T \leq 1$ the standard turbulence parameter of accretion disks. This Reynolds number is the only free parameter involved in the equations. The solutions shown in Fig. 9.15 ($a_H = 0.998$ and $\mathcal{R}_m = 25$) were calculated from an initial dipolar disk configuration based on advection and diffusion alone. After many diffusion times, a closed quadrupole for the toroidal field evolves around the horizon, which extends beyond the ergosphere. These poloidal current loops are mainly driven by the differential rotation of absolute space, $\nabla\omega$. They escape into an open structure at distances of a few gravitational radii. In the region, where these current loops cross the poloidal field lines, Lorentz forces $\mathbf{j}_p \times \mathbf{B}_p$ can accelerate plasma filled into the magnetosphere. This is the basic process which leads to high plasma acceleration in open magnetospheres around rapidly rotating black holes.

Quadrupolar magnetic configurations are the naturally excited magnetic structures in geometrically thin accretion disks. Disk outflows driven by quadrupolar structures cannot be efficiently accelerated, leading probably only to moderate disk winds, such as observed in the microquasars (typical speed is $0.9c$). For efficient plasma acceleration with Lorentz factors $\Gamma \simeq 10\text{--}20$, as observed in quasars, dipolar structures are needed, which, for some reasons, are preferably built up in elliptical galaxies.

9.5.4 Stationary Magnetospheres on Kerr Black Holes

The theory of stationary and axisymmetric MHD flows is now complete and can be used for any investigation of such flows in the gravitational field of rotating compact objects. The basic features have been derived by the present author [96, 97]. A recent review can be found in [98, 100, 101]. Here, we just make a few comments on the jet formation process. These developments will be superseeded by time-dependent MHD simulations which will be discussed in Sect. 10.5.

Force-Free Magnetospheres

As we have demonstrated in Fig. 9.15, the differential rotation of absolute space, $\nabla\omega$, produces current loops over the current flux equation (9.79). In this way, a black hole will be embedded into a rotating magnetosphere filled up with currents. This reminds us of the pulsar problem.

Goldreich and Julian [176] analyzed, already in 1970, the vacuum solution for a rotating neutron star with a dipolar magnetic field aligned with the rotational axis. They argued that the rotationally induced electric field was strong enough to pull charged particles from the stellar surface and, thus, fill the surrounding space with plasma. Using the force-free approximation to describe the produced magnetosphere they argued that an electromagnetically driven wind would carry away rotational energy and angular momentum of the star.

Wald found in 1974 [409] a rather interesting particular solution of the vacuum Maxwell equations in the Kerr spacetime. Far away from the hole this solution described a uniform magnetic field aligned with the rotational axis of a black hole. However, near the black hole it described a strong electric field as well.

Finally, Blandford and Znajek [80] realized in 1977 that the similarity between the vacuum solution for a Kerr black hole and the vacuum solution for a rotating neutron star meant the possibility of electromagnetically driven wind from a rotating black hole, provided the space around the black hole could be filled with plasma. Moreover, they argued that, under the typical astrophysical conditions, the vacuum solutions were, in fact, unstable to cascade pair production, ensuring a plentiful supply of charged particles. Then they developed a general theory of force-free steady-state axisymmetric magnetospheres of black holes and found a perturbative solution for a slowly rotating black hole with monopole magnetic field. The key element of this solution was Znajek's “boundary condition” imposed on the event horizon. As expected, this solution exhibited outgoing electromagnetic fluxes of energy and angular momentum. Moreover, the electromagnetic mechanism seemed to be very robust and the estimated power of the wind was high enough to explain the energetics of radio galaxies and quasars.

The electrodynamic mechanism together with the horizon theory is now widely accepted by the astrophysical community. In great contrast to this mainstream trend, Punsly and Coroniti [335, 336] and later Punsly (see [337]) completely rejected both these theories. They argued that the event horizon cannot be regarded as a unipolar inductor because it is causally disconnected from the outgoing wind. Indeed, both the

fast and the Alfvén waves generated at the event horizon can propagate only inwards and cannot effect the events in the outer space. The apparent lack of a proper unipolar inductor in the Blandford–Znajek solution and its reliance on Znajek’s boundary condition made Punsly and Coroniti to conclude that this solution is nonphysical and structurally unstable. They developed completely different MHD models which seemed to be based on clearer physical ideas [337]. In brief, they argue that gravity forces magnetospheric plasma to rotate inside the black hole ergosphere in the same sense as the black hole and that the magnetic field exhibits a similar rotation because it is “frozen” into this plasma.

Advection of magnetic flux by the accretion process will immerse the black hole into an axisymmetric magnetosphere with a flux distribution given by

$$\Psi_d(r, \theta) = \frac{\Psi_\infty}{2\sqrt{\pi\Gamma(r)}} \exp\left(-\frac{\xi^2}{4\Gamma(r)}\right). \quad (9.83)$$

$\Gamma(r)$ is a slowly varying covariance function for the Gaussian flux distribution along spherical surfaces $\xi = (\pi/2 - \theta)/\Theta_D$. Ψ_∞ is a measure for the asymptotic flux advected inwards along the disk. In elliptical galaxies we expect $\Psi_0 \simeq 10^{33}$ gauss cm², corresponding to field strengths of the order of a milligauss on the parsec-scale. Near the horizon, the covariance function stays nearly constant so that the total flux advected inwards towards the horizon is

$$\Psi_H = \frac{\Psi_0}{2\sqrt{\pi\Gamma(r_H)}}. \quad (9.84)$$

Near the horizon, the field in the disk is practically only radial, since $\Psi_d = \Psi_d(\theta)$. It decays as a Gaussian distribution in vertical direction. Beyond the scale-height Θ_D , depending on the accretion process, this magnetosphere is largely force-free since it is filled up with a sparse plasma density. In the disk itself, the magnetic structure is far from being force-free, turbulent conductivity is necessary to provide the diffusion of the plasma against magnetic fields. Electromagnetic fields in the region between the disk surface and the rotational axis satisfy therefore the force-free condition (Fig. 9.19)

$$\varrho_e \mathbf{E}_\perp + \frac{1}{c} \mathbf{j} \times \mathbf{B} \simeq 0. \quad (9.85)$$

The Grad–Shafranov Equation

The structure of the resulting magnetosphere is then a solution of the stationary and axisymmetric Maxwell’s equations

$$\nabla \cdot \mathbf{E} = 4\pi\varrho_e, \quad \nabla \cdot \mathbf{B} = 0 \quad (9.86)$$

$$\nabla \times (\alpha \mathbf{E}) = (\mathbf{B} \cdot \nabla \omega) \mathbf{m} \quad (9.87)$$

$$\nabla \times (\alpha \mathbf{B}) = 4\pi\alpha \mathbf{j} - (\mathbf{E} \cdot \nabla \omega) \mathbf{m}. \quad (9.88)$$

Under axisymmetry we decompose the fields into poloidal and toroidal components, $\mathbf{B} = \mathbf{B}_p + \mathbf{B}_T$. The force-free condition (9.85) then implies $\mathbf{E} \cdot \mathbf{B} = 0$, i.e. parallel electric fields must vanish. Charged particles are forced to flow along magnetic field lines. Hence, the angular velocity of the field lines Ω_F is given relative to the ZAMO by

$$\mathbf{v}_F = \frac{\Omega_F - \omega}{\alpha} \tilde{\omega} \mathbf{e}_\phi = \frac{\Omega_F - \omega}{\alpha} \mathbf{m}. \quad (9.89)$$

The toroidal current density is therefore $\mathbf{j}_T = \mathbf{v}_F \varrho_e$. The force-free condition (9.85) then implies

$$\mathbf{E} = \mathbf{E}_\perp = -\frac{1}{\varrho_e} \mathbf{j}_T \times \mathbf{B}_p. \quad (9.90)$$

The poloidal magnetic field can be expressed in terms of the magnetic flux Ψ covered by a rotationally symmetric flux surface A

$$\Psi = \int_A \mathbf{B} \cdot d\mathbf{S}. \quad (9.91)$$

This means that

$$\boxed{\mathbf{B}_p = \frac{\nabla \Psi \times \mathbf{e}_\phi}{2\pi \tilde{\omega}} = \frac{\nabla \Psi \times \mathbf{m}}{2\pi \tilde{\omega}^2}}, \quad (9.92)$$

where we used $\mathbf{m} \cdot \mathbf{m} = g_{\phi\phi} = \tilde{\omega}^2$. Hence

$$\mathbf{E}_p = -\mathbf{v}_F \times \mathbf{B}_p = -\frac{\Omega_F - \omega}{2\pi\alpha} \nabla \Psi. \quad (9.93)$$

We can also consider the currents enclosed by the surface A

$$I = - \int_A \alpha \mathbf{j} \cdot d\mathbf{S}. \quad (9.94)$$

Then, similarly to the magnetic flux, $\nabla I = -2\pi \tilde{\omega} \mathbf{e}_\phi \times (\alpha \mathbf{j}_p)$, we obtain

$$\alpha \mathbf{j}_p = -\frac{\nabla I \times \mathbf{e}_\phi}{2\pi \tilde{\omega}} = -\frac{\nabla I \times \mathbf{m}}{2\pi \tilde{\omega}^2}. \quad (9.95)$$

Since $I = I(\Psi)$, we finally get the expression for the poloidal current density

$$\boxed{\mathbf{j}_p = -\frac{1}{\alpha} \frac{dI}{d\Psi} \mathbf{B}_p.} \quad (9.96)$$

Next, we consider Ampère's law

$$\begin{aligned} & \int_A \nabla \times (\alpha \mathbf{B}) \cdot d\mathbf{S} \\ &= 4\pi \int_A \alpha \mathbf{j} \cdot d\mathbf{S} - \int_A (\mathbf{E} \cdot \nabla \omega) \mathbf{m} \cdot d\mathbf{S}. \end{aligned} \quad (9.97)$$

Upon using Stokes' theorem and $\mathbf{m} \cdot d\mathbf{S} = \mathbf{m} \cdot (d\mathbf{r} \times 2\pi\mathbf{m}) = 0$, we derive the equality

$$\mathbf{B}_T = -\frac{2I}{\alpha\tilde{\omega}} \mathbf{e}_\phi = -\frac{2I}{\alpha\tilde{\omega}^2} \mathbf{m}. \quad (9.98)$$

Finally, we use Gauss' law to derive an expression for the charge density

$$\varrho_e = \frac{1}{4\pi} \nabla \cdot \mathbf{E}_\perp = -\frac{1}{8\pi^2} \nabla \cdot \left[\frac{\Omega_F - \omega}{\alpha} \nabla \Psi \right], \quad (9.99)$$

while Ampère's law yields

$$\begin{aligned} \mathbf{j}_T &= \frac{1}{4\pi\alpha} \{ [\nabla \times \mathbf{B}]_T + \tilde{\omega}(\mathbf{E} \cdot \nabla \omega) \} \\ &= -\frac{\tilde{\omega}}{8\pi^2\alpha} \left[\nabla \cdot \left(\frac{\alpha}{\tilde{\omega}^2} \nabla \Psi \right) + \frac{\Omega_F - \omega}{\alpha} \nabla \Psi \cdot \nabla \omega \right]. \end{aligned} \quad (9.100)$$

We now consider the poloidal component of the force-free condition

$$-\varrho_e(\mathbf{v}_F \times \mathbf{B})_p + (\mathbf{j} \times \mathbf{B})_p = 0. \quad (9.101)$$

This is equivalent to

$$-\varrho_e \frac{\Omega_F - \omega}{\alpha} \tilde{\omega} + j_T + \frac{1}{\alpha} \frac{dI}{d\Psi} B_T = 0. \quad (9.102)$$

By replacing the two expressions for ϱ_e and j_T , we obtain the famous **Grad-Shafranov equation** (GSS) for relativistic magnetospheres

$$\nabla \cdot \left[\frac{\alpha D}{\tilde{\omega}^2} \nabla \Psi \right] + \frac{\Omega_F - \omega}{\alpha} \frac{d\Omega_F}{d\Psi} |\nabla \Psi|^2 + \frac{16\pi^2 I}{\alpha \tilde{\omega}^2} \frac{dI}{d\Psi} = 0.$$

(9.103)

The structure of the magnetosphere depends on the field line rotation law $\Omega_F(\Psi)$ and the current distribution $I(\Psi)$ in the magnetosphere. In total, the GSS equation is a highly nonlinear partial differential equation, and only very special analytic solutions have been found, essentially for a linear current function, $I(\Psi) \propto \Psi$.

The Role of the Light Cylinder Surfaces

The function D is defined as

$$D = 1 - \frac{(\Omega_F - \omega)^2 \tilde{\omega}^2}{c^2 \alpha^2}$$

(9.104)

and is called the *light cylinder function*. In the special relativistic case, where $\omega = 0$ and $\alpha = 1$, $D = 0$ defines the light cylinder surface with the light cylinder radius

given by $R_L = c/\Omega_F$. In Kerr, the light cylinder surfaces are located at the two positions

$$\tilde{\omega}_{\text{ILC}} = \frac{\alpha c}{\omega - \Omega_F}, \quad \tilde{\omega}_{\text{OLC}} = \frac{\alpha c}{\Omega_F - \omega}. \quad (9.105)$$

They are called inner light cylinder surface (ILC) and outer light cylinder surface. This is the greatest distinction between a pulsar magnetosphere, which only has the outer light cylinder surface, and black hole magnetospheres.

The GSS equation has a singular behavior at the light cylinder, or in other words, the light cylinder surface is a critical surface for the GSS equation. Besides the light cylinder function, the structure of the resulting magnetosphere is determined by the current system flowing in the magnetosphere, $I = I(\Psi)$, and by the differential rotation of the field lines. When field lines are anchored into a disk, their rotation is determined by the disk rotation (e.g. Keplerian). For field lines entering the horizon, their rotation is a priori not determined.

Since the light cylinder function depends quadratically on the cylinder radius, a magnetosphere around black holes has two light cylinder surfaces, given by the condition

$$D = 1 - \frac{(\Omega_F - \omega)^2 \varpi_L^2}{c^2 \alpha^2} = 0 \quad (9.106)$$

for given rotation Ω_F . The outer light cylinder surface (OLC) replaces the special relativistic one and is slightly modified by gravitational effects near the equator. In addition, we find an inner light cylinder surface (ILC) which always is located within the ergosphere (Fig. 9.16).

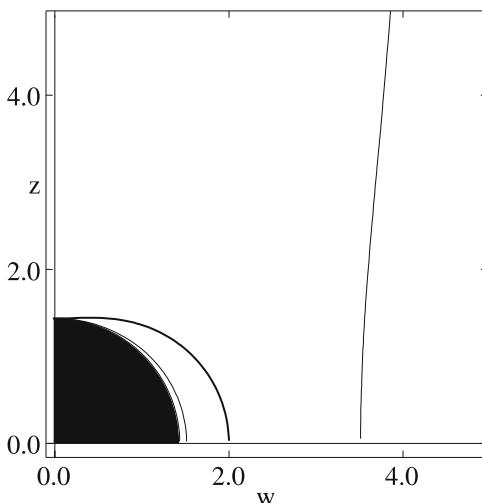


Fig. 9.16. Light cylinder surfaces in Kerr for a constant Ω_F . Coordinates are given in units of gravitational radii, GM/c^2 . The *thick line* denotes the ergosphere, the *inner light cylinder* is always between the horizon and the static limit surface. The *outer light cylinder* is slightly deformed by gravity near the equatorial plane

We can get closed expressions for the charge density and the toroidal current density [314]

$$\varrho_e = \frac{\Omega_F - \omega}{4\pi\alpha^2} \frac{(8\pi^2 I/\alpha^2)(dI/d\Psi) - \mathbf{N} \cdot \nabla\Psi}{D} \quad (9.107)$$

$$j_T = \frac{1}{4\pi^2\varpi} \frac{(8\pi^2 I/\alpha^2)(dI/d\Psi) - [(\Omega_F - \omega)^2 \varpi^2/\alpha^2] \mathbf{N} \cdot \nabla\Psi}{D}, \quad (9.108)$$

where

$$\mathbf{N} = \frac{1}{2} \left[\nabla \ln \frac{(\Omega_F - \omega)\varpi^2}{\alpha^2} - \frac{(\Omega_F - \omega)\varpi^2}{\alpha^2} \nabla\omega \right]. \quad (9.109)$$

On the light cylinder surfaces, where $D = 0$, the numerators in these expressions should also vanish in order to have a finite charge density and toroidal current density. This requires a boundary condition

$$\mathbf{N} \cdot \nabla\Psi = \frac{4\pi}{\alpha^2} \left(\frac{dI^2}{d\Psi} \right)_{LC}, \quad (9.110)$$

which is called a critical condition. According to this, the magnetosphere of a black hole subdivides into three regions: (i) the magnetosphere beyond the outer light cylinder surface, (ii) the region between the inner and outer light cylinder surface, and (iii) the magnetosphere between the horizon and the inner light cylinder surface.

As was first pointed out by Goldreich and Julian, the original source of the plasma that will fill the magnetosphere of a pulsar is the surface of the pulsar itself. Strong electric fields on the pulsar surface rip off charged particles from the star, and they fill the whole magnetosphere. In fact, this charge density is only a tiny fraction of the total population of particles of the magnetosphere which consist of electron positron pairs created in the polar cap region.

In the case of rotating black holes, the source of the charged particles that will populate the magnetosphere is rather different. As one moves from the inner light cylinder toward the outer one, the angular velocity of the field lines grows from $\Omega_F < \omega$ to $\Omega_F > \omega$ (Fig. 9.17). Thus somewhere between the two light surfaces, there should be a *null surface*, where $\Omega_F = \omega$, i.e. $\mathbf{v}_F = 0$. On this surface, the electric field will vanish, $\mathbf{E}_\perp = 0$ (in Fig. 9.17 this occurs near the radius of the marginal stable orbit). This null surface could be a spark gap, or a particle creation zone for nearly neutral particles. This could also be the zone, where plasma from the surrounding disk is injected. Plasma outside the null surface will be blown away by centrifugal forces, plasma inside the null surface will accrete towards the horizon. The null zone subdivides the magnetosphere of a black hole into the accretion zone near the horizon, and the wind zone beyond the null surface.

The pulsar equation is a special example of the GSS equation in flat Minkowski space, i.e. for $\alpha = 1$ and $\omega = 0$. Numerical solutions for a dipolar magnetosphere of rotating neutron stars have been constructed by Contopoulos et al. [122], and recently by Timokhin [390] (Fig. 9.18).

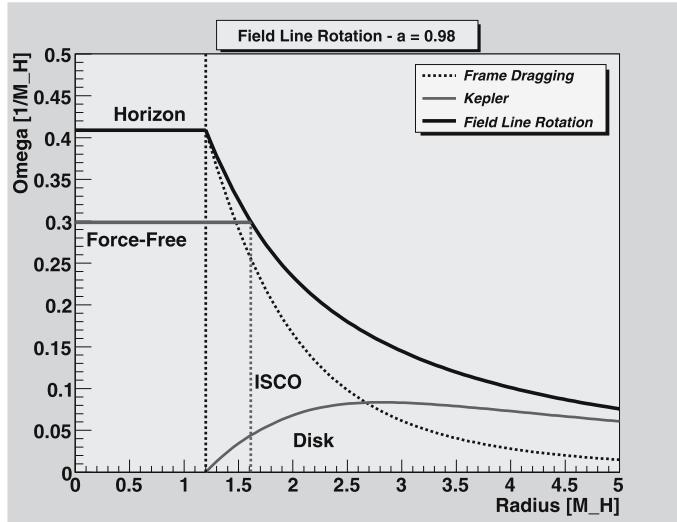


Fig. 9.17. Field line rotation of the magnetosphere around a rapidly rotating black hole ($a = 0.98$). The *dashed vertical line* gives the position of the horizon. The *solid curve* shows the rotation of field lines anchored in the disk. Near the horizon, maximum rotation is not reached, so that $\Omega_F < \Omega_H$ (here given for $\Omega_F \simeq \Omega(r_{\text{ms}})$). Note that within three gravitational radii, field line rotation around a rapidly rotating black hole is completely dominated by frame-dragging (dotted curve)

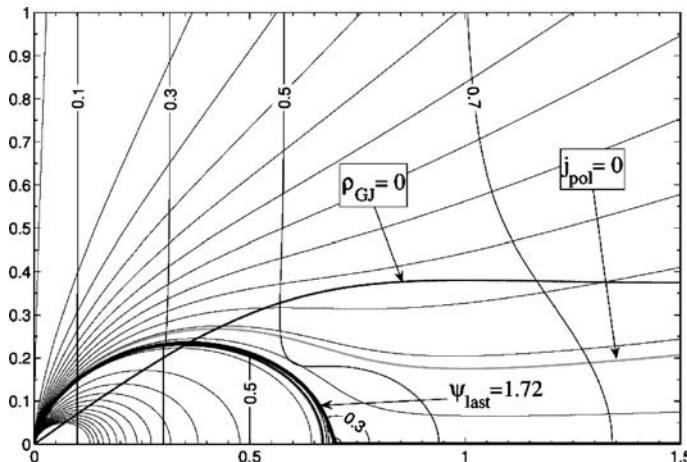


Fig. 9.18. Axisymmetric magnetosphere of rotating neutron stars (the Goldreich–Julian problem). Coordinates are given in units of light cylinder radii. Part of the dipolar magnetosphere is closed within the light cylinder. An equatorial current sheet with the return current is included. The equatorial current sheet can be excluded from the numerical treatment by setting appropriate boundary conditions at the equatorial plane. Figure adapted from [390]

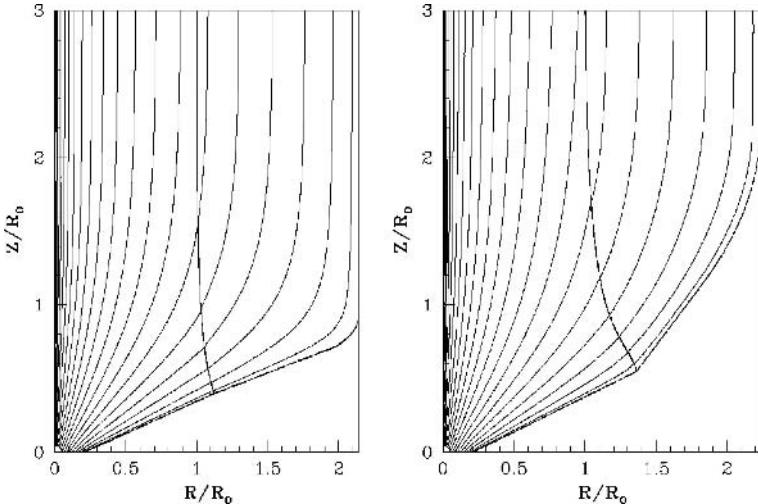


Fig. 9.19. The distribution of magnetic flux in the meridional plane of a black hole. A finite accretion disk carries some dipolar magnetic flux which is stretched by a disk wind into the upper hemisphere. In this approximation, the magnetosphere is assumed to be force-free for simplification. The contour levels of the magnetic flux are given by $\Psi = 10^a$ with $a = 0.1n^2$. R_0 is the asymptotic light cylinder, while the light cylinder increases for flux surfaces anchored at larger disk radii due to differential rotation in the disk (vertical line). Figures adapted from Fendt and Memola [157]

Fendt [156, 157] discussed solutions of the GSS equation obtained in a general relativistic context applying the 3+1 formalism for Kerr geometry. These solutions extend from the inner light surface around the Kerr black hole to the asymptotic regime of a cylindrically collimated jet with a finite radius (Fig. 9.19).

As asymptotic boundary condition, the analytical special relativistic 1D GSS solution of Appl and Camenzind [37] is implemented. Since $I = I(\Psi)$, this asymptotic solution provides the GSS source term also for the collimation region. The disk magnetic field distribution is parameterized as, $x = R/R_L$, $\Psi \sim (x/b)^m / (1 + (x/b)^m) + \Psi_H$. Here, the magnetic flux covered by the black hole is $\Psi_H = 0.5$ with $m = 3$ and a core radius b . The asymptotic jet radius can be parameterized in terms of the light cylinder radius R_L or the gravitational radius M . For $a = 0.8$, $(\Omega_F/\Omega_H) = 0.4$, the jet radius is $3 R_L$ corresponding to $30 M$. The main features of the calculated jet magnetosphere are the following (Fig. 9.19, left). The field lines originate near the inner light surface close to the rotating black hole and collimate to an asymptotic jet of finite radius of several (asymptotic) light cylinder radii. The solution is defined on a global scale, satisfying the regularity condition (9.110) along the light surfaces. A rapid field collimation occurs within $20 M$ distance from the source. The near-disk solution has three different regimes. Here, the magnetic flux is either outgoing towards the asymptotic jet or in-going towards the black hole. But there exist also flux surfaces near the axis which are not connected to the disk.

The field structure is governed by $I(\Psi)$ and $\Omega_F(\Psi)$. In combination with the disk magnetic flux distribution, this allows us to determine the magnetic angular momentum loss from the disk in the jet and the toroidal magnetic field along the disk. The angular momentum flux per unit time per unit radius is $d\dot{J}/dx = -xB_z I(x)$ along the disk. In this solutions, most of the magnetic angular momentum is lost in the outer part of the disk.

Beyond the Force-Free Approximation

In a black hole magnetosphere, because of the strong gravity of the black hole and the rapid rotation of the magnetic field, both an ingoing plasma flow (accretion) and an accelerated outgoing plasma (wind/jet) should be created. The plasma would be provided from the disk surface and its corona. When the plasma density in the magnetosphere is somewhat large, the plasma inertia effects should be important. In this case, the plasma would be nearly neutral and should be treated by the ideal magnetohydrodynamic (MHD) approximation, so the plasma streams along a magnetic field line, where the magnetic field line could extend from the disk surface to the event horizon or a far distant region [96]. The outgoing flow effectively carries the angular momentum from the plasma source, and then the accretion would continue to be stationary, releasing its gravitational energy. The magnetic field lines connecting the black hole with the disk, which are mainly generated by the disk current, may not connect directly to the distant region, but via the disk's interior the energy and angular momentum of the black hole can be carried to the distant region; the energy and angular momentum transport inside the disk is not discussed here.

We treat a stationary and axisymmetric magnetosphere, and consider ideal MHD flows along a magnetic field line. The initial velocity can be at most less than the slow magnetosonic wave speed. To accrete onto the black hole, the ejected inflows from the plasma source must pass through the slow magnetosonic point (**S**), the Alfvén point (**A**) and the fast magnetosonic point (**F**) in this order. At these points, **A**, **F** and **S**, the poloidal velocity equals one of the Alfvén wave and fast and slow magnetosonic wave speeds, respectively. In the case of accretion onto a star, because the accreting plasma is stopped at the stellar surface, a shock front would be formed somewhere on the way to the stellar surface and the accretion becomes subfast magnetosonic. However, for accretion onto a black hole, the flow must be superfast magnetosonic at the event horizon (**H**). If not so, the fast magnetosonic wave can extract information from the interior of the black hole to the exterior; this fact obviously contradicts with the definition of the event horizon. In fact, an ideal MHD accretion solution which keeps subfast magnetosonic has zero poloidal velocity at the event horizon and the density of the plasma diverges – this solution would be unphysical.

Because the magnetic field lines would rigidly rotate under the ideal MHD assumption, there are two light surfaces (**LC**) in the black hole magnetosphere. The plasma source must be located between these two surfaces. Further, one or two Alfvén surfaces lie between the two light surfaces, and for accretion there must be

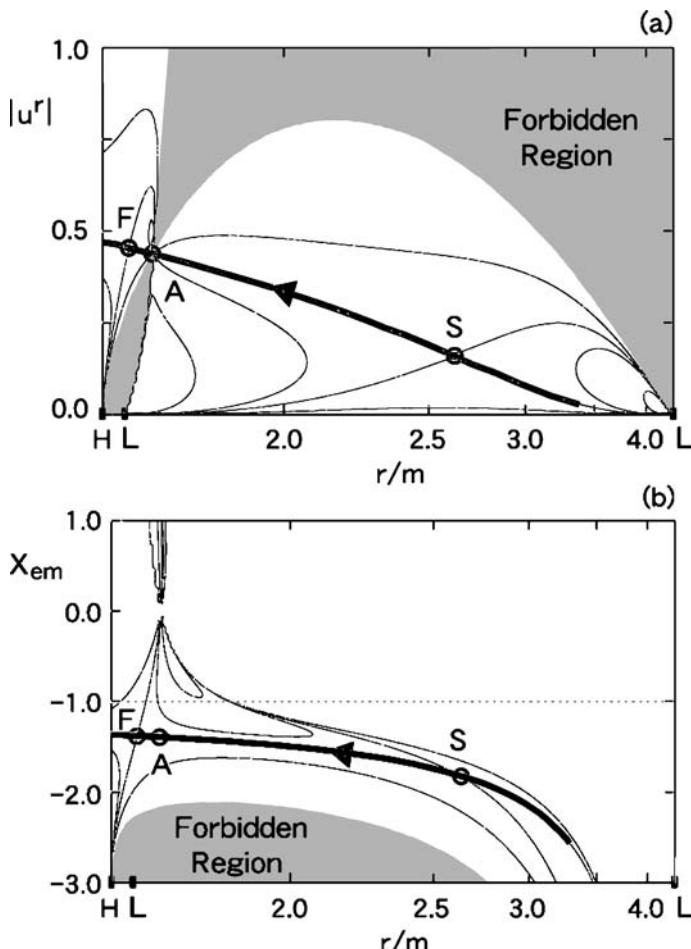


Fig. 9.20. A negative energy trans-magnetosonic MHD accretion solution in the Kerr geometry (*thick curves*, $a = 0.8$ and $\Omega_F = 0.6\Omega_H$). *Top panel:* radial four-velocity; *lower panel:* the ratio of the electromagnetic energy to the total energy, $X_{em} = -\Omega_F B_\phi / 4\pi\eta|E|$. The accretion flow starts near the marginal stable orbit, passes the slow magnetosonic point (S) and the Alfvén point (A) within the ergosphere, and goes through the fast magnetosonic point (F) just outside the horizon. Figure adapted from [385]

a fast-magnetosonic surface between the Alfvén surface and the event horizon. Here, we should note that the physical mechanism to determine the angular velocity of the field lines is controversial. A time-dependent determination of it has been discussed; the torsional Alfvén wave originated from the plasma source and propagated up and down the magnetic flux tube forces to minimize the magnetic stresses in the system.

The stationary relativistic hot MHD flow equations have been formulated by [96–98]. For the details see Problem 9.8. Stationary and axisymmetric plasma

flows are determined by five integrals of motion (constants on magnetic surfaces)

$$\{\Omega_F(\Psi), \eta(\Psi), s(\Psi), E(\Psi), L(\Psi)\}. \quad (9.111)$$

$\Omega_F(\Psi)$ is the angular velocity of field lines, as in the force-free case, $\eta(\Psi)$ denotes the plasma load per magnetic flux tube, $E(\Psi)$ and $L(\Psi)$ are the total energy and angular momentum, respectively, carried by the plasma flow, and $s(\Psi)$ is the entropy per flux tube, which is constant in adiabatic plasma flows. By using these conserved quantities, the equation of motion projected onto the direction of a poloidal magnetic field, which is called the poloidal equation (and is often referred to as the relativistic Bernoulli or wind equation), can be expressed as follows [385]

$$1 + u_p^2 = (E/\mu)^2 \left[(\alpha_g - 2M^2) f^2 - k \right]. \quad (9.112)$$

Here we use the following abbreviations

$$\alpha_g = g_{tt} + 2g_{t\phi}\Omega_F + g_{\phi\phi}\Omega_F^2 \quad (9.113)$$

$$k = (g_{\phi\phi} + 2g_{t\phi}\tilde{L} + g_{tt}\tilde{L}^2)/\varrho_w^2 \quad (9.114)$$

$$f = -\frac{(g_{t\phi} + g_{\phi\phi}\Omega_F) + (g_{tt} + g_{t\phi}\Omega_F)\tilde{L}}{\varrho_w(M^2 - \alpha_g)}, \quad (9.115)$$

where $\tilde{L} = L/E$ and $\varrho_w^2 \equiv g_{t\phi}^2 - g_{tt}g_{\phi\phi}$. In addition, we use $\mathcal{R} = \tilde{e}^2\alpha_g - 2\tilde{e}^2M^2 - kM^4$ and $\tilde{e} = 1 - \Omega_F L/E$.

The relativistic Alfvén Mach number M is defined by

$$M^2 \equiv \frac{4\pi\mu n u_p^2}{B_p^2} = \frac{4\pi\mu\eta u_p}{B_p}. \quad (9.116)$$

The wind equation is essentially an equation for the Mach number along the flux tube

$$M^2 = M^2(r; \mathbf{B}_p^2, \Omega_F(\Psi), \eta(\Psi), E(\Psi), L(\Psi), s(\Psi)). \quad (9.117)$$

The locations of the Alfvén points (r_A, θ_A) along a magnetic field line, where $\theta = \theta(r; \Psi)$, are defined by $M^2 = \alpha_g$. A solution for the Mach number is shown in Fig. 9.21.

The differential form of the poloidal equation (9.112) is written as

$$(\ln u_p)' = \frac{\mathcal{N}}{\mathcal{D}},$$

(9.118)

where

$$\begin{aligned} \mathcal{N} &= \left(\frac{E}{\mu}\right)^2 \left\{ [\mathcal{R}(M^2 - \alpha_g)C_{sw}^2 + M^4\mathcal{A}^2] (\ln B_p)' \right. \\ &\quad \left. + \frac{1}{2}(1 + C_{sw}^2)[M^4(M^2 - \alpha_g)k' - \mathcal{Q}\alpha'_g] \right\} \end{aligned} \quad (9.119)$$

$$\mathcal{D} = (M^2 - \alpha_g)^2 [(C_{sw}^2 - u_p^2)(M^2 - \alpha_g) + (1 + u_p^2)M^4\mathcal{A}^2\mathcal{R}^{-1}]. \quad (9.120)$$

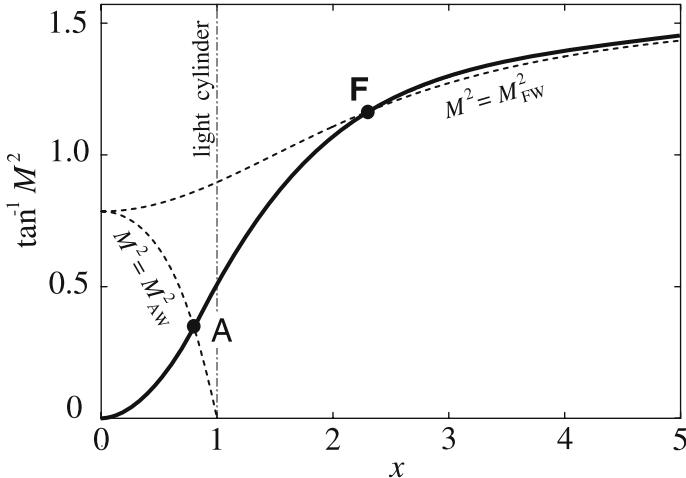


Fig. 9.21. Solution for a cold wind in a collimating magnetosphere as a function of radii in units of light cylinder radii, $x = R/R_L$. The solution passes smoothly through the Alfvén point (A) inside the light cylinder and the fast magnetosonic point (F) outside the light cylinder. Figure adapted from [385]

Here we used the quantities $\mathcal{A}^2 = \tilde{e}^2 + \alpha_g k = f^2(M^2 - \alpha_g)^2$ and $\mathcal{Q} = \alpha_g \tilde{e}^2 - 3\tilde{e}^2 M^2 - 2kM^4$. The prime $(\dots)'$ denotes $[(\partial_\theta \Psi)\partial_r - (\partial_r \Psi)\partial_\theta]/(\sqrt{-g}B_p)$ which is a derivative along a stream line. The relativistic sound velocity a_{sw} is given by

$$a_{\text{sw}}^2 \equiv \left(\frac{\partial \ln \mu}{\partial \ln n} \right)_{\text{ad}} = (\Gamma - 1) \frac{\mu - m_p}{\mu}, \quad (9.121)$$

and the sound four-velocity is given by $C_{\text{sw}}^2 = a_{\text{sw}}^2/(1 - a_{\text{sw}}^2)$.

The denominator in the above equation can be reduced to the form

$$\mathcal{D} \propto (u_p^2 - u_{\text{AW}}^2)^2 (u_p^2 - u_{\text{FM}}^2) (u_p^2 - u_{\text{SM}}^2), \quad (9.122)$$

where the relativistic Alfvén wave speed u_{AW} , the fast magnetosonic wave speed u_{FM} and the slow magnetosonic wave speed u_{SM} are defined by

$$u_{\text{AW}}^2(r; \Psi) \equiv \frac{B_p^2}{4\pi\mu n} \alpha \quad (9.123)$$

$$u_{\text{FM}}^2(r; \Psi) \equiv \frac{1}{2} \left(Z + \sqrt{Z^2 - 4C_{\text{sw}}^2 u_{\text{AW}}^2} \right) \quad (9.124)$$

$$u_{\text{SM}}^2(r; \Psi) \equiv \frac{1}{2} \left(Z - \sqrt{Z^2 - 4C_{\text{sw}}^2 u_{\text{AW}}^2} \right) \quad (9.125)$$

with

$$Z \equiv u_{\text{AW}}^2 + \frac{B_\phi^2}{4\pi\mu n Q_w^2} + C_{\text{sw}}^2, \quad (9.126)$$

$$B_\phi = -4\pi\eta E \varrho_w f. \quad (9.127)$$

When $u_p^2 = u_{\text{AW}}^2$, $u_p^2 = u_{\text{FM}}^2$ or $u_p^2 = u_{\text{SM}}^2$, the denominator in equation (9.118) becomes zero. Therefore, at these singular points, we must require $\mathcal{N} = 0$ to obtain physical accretion solutions which pass through these points smoothly. The location of $u_p^2 = u_A^2$ [$\equiv u_{\text{AW}}^2(r_A; \Psi)$] is the Alfvén point discussed in the previous section. Similarly, the locations of $u_p^2 = u_F^2$ [$\equiv u_{\text{FM}}^2(r_F; \Psi)$] and $u_p^2 = u_S^2$ [$\equiv u_{\text{SM}}^2(r_S; \Psi)$] correspond to the fast magnetosonic point $r = r_F$ and the slow magnetosonic point $r = r_S$, respectively. We should mention that, to calculate the Alfvén velocity, we need to solve a polynomial of high degree, while in the cold limit it is simply obtained as $u_A^{\text{cold}} = (\alpha B_p)_A / (4\pi\mu_c\eta)$; the Alfvén velocity of a hot MHD flow is always smaller than u_A^{cold} .

Figure 9.20 shows a negative energy accretion solution (for $\Omega_F L/E > 1$). We see that the Alfvén point is located inside the ergosphere. The outgoing electromagnetic energy flux is always greater than the ingoing plasma energy flux ($X_{\text{em}} < -1$ and $X_{\text{plasma}} = -1 - X_{\text{em}} > 0$). The magnetic field lines are trailed ($B^\phi/B_p < 0$) everywhere due to the black hole rotation. **A Poynting flux flows away from the horizon into the disk plasma.**

For accretion with $\Omega_F L/E < 1$ and $0 < \Omega_F < \Omega_H$, the electromagnetic energy flux also streams outward everywhere, but at least near the event horizon the ingoing plasma energy flux dominates (i.e. $-1 < X_{\text{em}}(r_H) < 0$). For magnetically

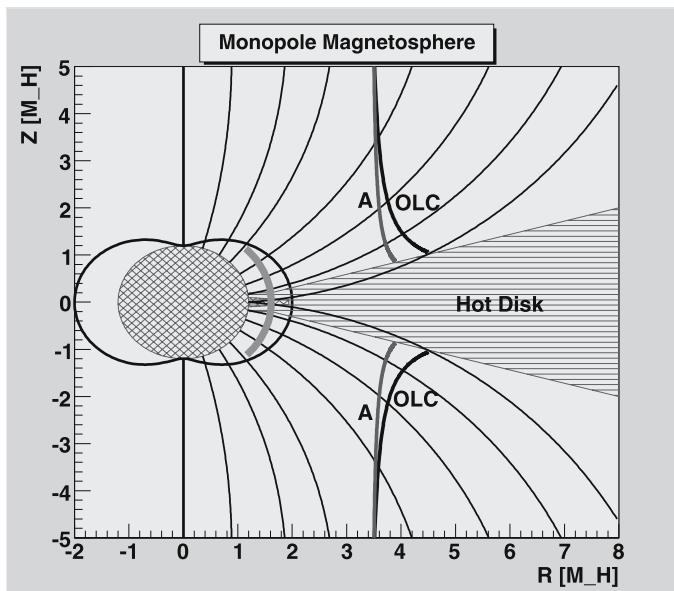


Fig. 9.22. Structure of the magnetosphere formed in accretion processes ($a = 0.98$). The *inner thick line* denotes the ergosphere, the *outer ones* the outer light cylinder surfaces (OLC) and the Alfvén surfaces (A). Near the horizon, a kind of neutral surface is formed in the equatorial plane. Plasma accretion occurs along radial flux tubes, and plasma is injected into the magnetosphere near the marginal stable orbit (*thick half circle*)

dominated accretion, we see that $X_{\text{em}} \simeq -1$. The Poynting flux passing through the event horizon is not modified by the plasma inertia effect. This is because the toroidal magnetic field at the event horizon becomes

$$B_{\phi,H} = \sqrt{(g_{\phi\phi}/\Sigma)_H} (\Omega_H - \Omega_F)(\partial_\theta\Psi)_H \quad (9.128)$$

for any ideal MHD accretion flows, which is the same expression as that of the force-free case (i.e. the Znajek horizon condition). This expression can be written in the form [72]

$$4\pi I(\Psi) = [\Omega_F - \Omega_H] \frac{M^2 + a^2}{M^2 + a^2 \cos^2 \theta} \sin \theta \left(\frac{d\Psi}{d\theta} \right)_H. \quad (9.129)$$

This condition is automatically true for any solution of the Grad–Shafranov equation, which can be extended up to the horizon.

The Blandford–Znajek process becomes more evident, when the finite inertia of plasma is included in the discussion. The total energy accreted by a black hole consists of two parts, the plasma part and the electromagnetic part (the latter is the only energy in the force-free approximation). Rapid rotation of magnetic fields near the horizon can produce then a negative Poynting flux, which is dominating for sufficiently strong magnetic fields [98]. Since the field lines are dragged radially toward the horizon (Fig. 9.22), the source term $\partial_r T \propto \alpha \varpi^2 (\mathbf{B}_p \cdot \nabla) \Omega$ in the current flux equation (9.79) is quite strong near the horizon. This effect will drive currents across the magnetosphere near the inner Alfvén surface, which is located near the inner light cylinder surface. When plasma conductivities are included in the treatment of the magnetospheric flows, it should be possible to show that the current system is closed not on the horizon, but within the ergosphere (see Fig. 9.15). The neutron star surface, which acts as the medium to close the magnetospheric current system in the Goldreich–Julian problem, is replaced in the case of rotating black holes by the ergospheric region. A method to treat time-dependent electrodynamics in the conservative formulation is discussed in the next section.

9.5.5 Relaxation of Black Hole Magnetospheres and the Blandford–Znajek Process

The boundary conditions at the horizon are crucial for understanding stationary magnetospheres around black holes. The horizon is however not a true boundary, but only a critical surface in the force-free approximation. It is identical with the inner light cylinder of the magnetosphere. Another technique to solve the magnetospheric problem is then to consider the time relaxation of a given initial magnetosphere by solving the time-dependent Maxwell's equations. This technique has been developed by Komissarov [238] quite recently.

Driven by the need to write the equations of black hole electrodynamics in a form convenient for numerical applications, Komissarov [238] has recently constructed a new system of 3+1 equations, which not only has a more traditional form than the now classic 3+1 system of Thorne and MacDonald, but also is more general. To deal

with the magnetospheric current sheets, he also developed a simple model of radiative resistivity based on the inverse Compton scattering of background photons. The results of numerical simulations combined with simple analytical arguments allow us to make a number of important conclusions on the nature of the Blandford–Znajek mechanism. Just like in the Penrose mechanism, the key role in this mechanism is played by the black hole ergosphere. The poloidal currents are driven by the gravitationally induced electric field, which cannot be screened within the ergosphere by any static distribution of the electric charge of locally created pair plasma. Contrary to what is expected in the membrane paradigm, the energy and angular momentum are extracted not only along the magnetic field lines penetrating the event horizon but also along all field lines penetrating the ergosphere. In dipolar magnetic configurations symmetric relative to the equatorial plane, the force-free approximation breaks down within the ergosphere, where a strong current sheet develops along the equatorial plane. This current sheet supplies energy and angular momentum at infinity to the surrounding force-free magnetosphere. The Blandford–Znajek monopole solution is found to be asymptotically stable and causal. The Znajek horizon boundary condition is shown to be a regularity condition at fast critical surface.

Maxwell's Equations in Conservative Form

The covariant Maxwell equations are

$$\nabla_\beta {}^*F^{\alpha\beta} = 0 \quad (9.130)$$

and

$$\nabla_\beta F^{\alpha\beta} = 4\pi I^\alpha, \quad (9.131)$$

where $F^{\alpha\beta}$ is the Maxwell tensor of the electromagnetic field, ${}^*F^{\alpha\beta}$ is the Faraday tensor and I^α is the four-vector of the electric current. The most direct way of 3+1 splitting of the covariant Maxwell equations is to write them down in components and then to introduce such spatial vectors that these equations have a particularly simple and familiar form. For example, when equation (9.130) is written in components it splits into two parts:

- The time part

$$\frac{1}{\sqrt{\gamma}} \partial_t (\alpha \sqrt{\gamma} {}^*F^{ti}) = 0. \quad (9.132)$$

- The spatial part

$$\frac{1}{\sqrt{\gamma}} \partial_i (\alpha \sqrt{\gamma} {}^*F^{ji}) + \frac{1}{\sqrt{\gamma}} \partial_t (\alpha \sqrt{\gamma} {}^*F^{ti}) = 0. \quad (9.133)$$

If we now introduce the spatial vectors \mathbf{B} and \mathbf{E} via

$$B^i = \alpha {}^*F^{it} \quad (9.134)$$

and

$$E^i = \gamma^{ij} E_j, \quad E_i = \frac{\alpha}{2} \eta_{ijk} {}^*F^{jk}, \quad (9.135)$$

where $\eta_{ijk} = \sqrt{\gamma} \epsilon_{ijk}$ is the Levi-Civita pseudotensor of the absolute space, then equations (9.132) and (9.133) read

$$\nabla \cdot \mathbf{B} = 0, \quad (9.136)$$

$$\partial_t \mathbf{B} + \nabla \times \mathbf{E} = 0, \quad (9.137)$$

where ∇ is the covariant derivative of the absolute space. Similarly, equation (9.131) splits into

$$\nabla \cdot \mathbf{D} = 4\pi \varrho_e, \quad (9.138)$$

$$-\partial_t \mathbf{D} + \nabla \times \mathbf{H} = 4\pi \mathbf{J}, \quad (9.139)$$

where

$$D^i = \alpha F^{ti} \quad (9.140)$$

$$H^i = \gamma^{ij} H_j, \quad H_i = \frac{\alpha}{2} \eta_{ijk} F^{jk} \quad (9.141)$$

and

$$\varrho_e = \alpha I^t, \quad J^k = \alpha I^k. \quad (9.142)$$

As one can see, these 3+1 equations have exactly the same form as the classical Maxwell equations for the electromagnetic field in matter. Applying ∇ to equation (9.139) and then using equation (9.138) one obtains the electric charge conservation law

$$\partial_t \varrho_e + \nabla \cdot \mathbf{J} = 0. \quad (9.143)$$

With the definition of these fields, Maxwell's equations in this formulation can be written in closed conservative form

$$\boxed{\partial_t \mathbf{U}_M + \partial_j [\sqrt{\gamma} \mathbf{F}_M^j] = \mathcal{S}_M,}$$

(9.144)

where \mathbf{U}_M is a seven-dimensional state vector consisting of the components $\mathbf{U}_M = (\Psi/\alpha, B^i + \Psi\beta^i/\alpha, D^i)^T$ and corresponding fluxes

$$\mathbf{F}_M^j = (B^j - \Psi\beta^j/\alpha, \eta^{ijk}E_k + \alpha\Psi g^{ij}, -\eta^{ijk}H_k)^T. \quad (9.145)$$

The corresponding source terms are as follows

$$\mathcal{S}_M = \left(-\frac{\Psi}{\sqrt{\gamma}}\partial_k(\sqrt{\gamma}\beta^k/\alpha), \frac{\Psi}{\sqrt{\gamma}}\partial_k(\sqrt{\gamma}\alpha g^{ik}), \varrho(-\alpha v_d^i + \beta^i) \right)^T. \quad (9.146)$$

Ψ is a pseudopotential introduced to clean the divergence of the magnetic fields, i.e. to numerically guarantee $\nabla \cdot \mathbf{B} = 0$ for all time-steps.

In Boyer–Lindquist coordinates, where $\nabla \cdot \boldsymbol{\beta} = 0$, these equations reduce to the corresponding equations we derived for the membrane formalism. For example, one finds that

$$\nabla \times \mathbf{E} = \nabla \times \alpha \mathbf{E} - \mathcal{L}_{\boldsymbol{\beta}} \mathbf{B}, \quad (9.147)$$

where

$$\mathcal{L}_{\boldsymbol{\beta}} \mathbf{B} = (\boldsymbol{\beta} \cdot \nabla) \mathbf{B} - (\mathbf{B} \cdot \nabla) \boldsymbol{\beta} \quad (9.148)$$

is the Lie derivative of \mathbf{B} along $\boldsymbol{\beta}$. Thus, equation (9.137) reads

$$\partial_t \mathbf{B} - \mathcal{L}_{\boldsymbol{\beta}} \mathbf{B} + \nabla \times (\alpha \mathbf{E}) = 0, \quad (9.149)$$

which is the standard induction equation in the membrane paradigm (see Sect. 8.6.1). However, in other coordinate systems, e.g. the Kerr–Schild system, $\nabla \cdot \boldsymbol{\beta} \neq 0$.

Maxwell's equations are completed by means of Ohm's law in the form

$$\mathbf{J} = \sigma_{\parallel} \mathbf{E}_{\parallel} + \sigma_{\perp} \mathbf{E}_{\perp} + \mathbf{j}_d, \quad (9.150)$$

where \mathbf{E}_{\parallel} is the electric field parallel to the magnetic field and \mathbf{E}_{\perp} the perpendicular component, and \mathbf{j}_d is a drift current which is perpendicular to both electric and magnetic fields. A particular model for the conductivities is discussed in [238]. The effect of the finite conductivity shows up in Fig. 9.23, where a split monopole has been relaxed.

The Relaxation of the Wald Magnetosphere

As an application of this technique we consider the relaxation of the Wald magnetosphere. Wald [409] obtained the following *vacuum solution* for a rotating black hole immersed into a uniform magnetic field aligned with the rotation axis

$$F_{\mu\nu} = B_{\infty} (m_{[\mu,\nu]} + 2ak_{[\mu,\nu]}), \quad (9.151)$$

where $k = \partial_t$ and $m = \partial_{\phi}$ are the Killing fields of the Kerr solution. Outside the horizon, this solution satisfies the vacuum field equations with

$$\mathbf{D} = -\frac{1}{\alpha} \boldsymbol{\beta} \times \mathbf{B}. \quad (9.152)$$

Thus $D_{\parallel} = 0$, and no electric current is driven along the magnetic field. Outside the horizon, $B^2 - D^2$ is strictly positive.

The Blandford–Znajek solution is the only global analytical solution for magnetospheres of rotating black holes found so far and for this reason it has been playing a key role in the development of the black hole electrodynamics. One important property of this solution is that all magnetic field lines penetrate the black hole horizon. The remaining magnetic flux splits between the field lines originating from the accretion disk and the field lines passing through the gap between the hole and the disk. In general, the angular velocity of magnetic field lines in steady-state force-free magnetospheres has to be prescribed, so one faces the task of setting physically sensible boundary conditions for all these three different types of magnetic field lines. In the case of the field lines originating from the accretion disk the solution is obvious. Their angular velocity is given by the angular velocity of the disk at the foot points. As for the other two kinds of magnetic field lines, this task is less trivial. In their solution, Macdonald and Thorne [263] and later Macdonald [262] appealed to the existing analogy between the black hole horizon and a rotating conducting sphere. They concluded that only the field lines penetrating the event horizon rotate, whereas in the gap $\Omega_F = 0$.

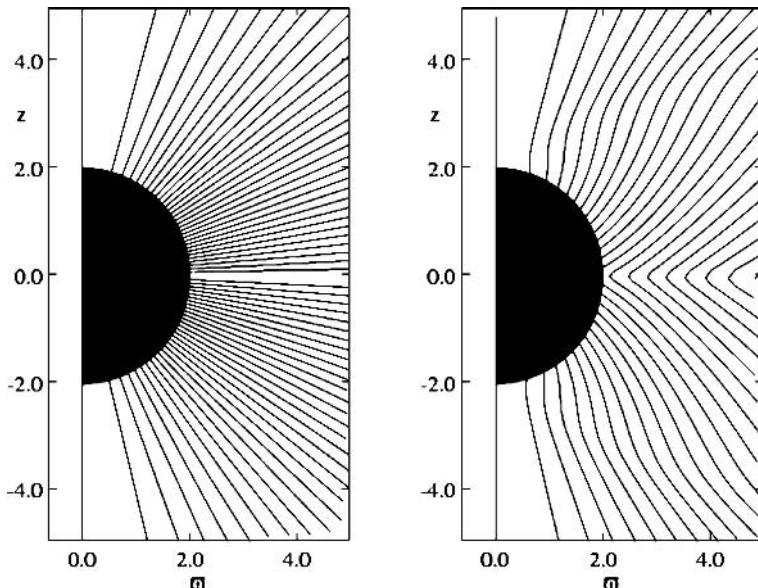


Fig. 9.23. Evolution of a split monopole around Schwarzschild black holes. *Left:* split monopole as initial condition; *right:* magnetic flux surfaces in the final evolution. Near the equatorial plane, a current sheet has developed. Figure adapted from Komissarov [238]

Figure 9.24 shows the final solution, at $t = 126$, for a Kerr black hole with $a = 0.9$ immersed into a Wald magnetosphere. A current sheet is formed in the equatorial plane within the black hole ergosphere. This is clearly seen in the right panel of Fig. 9.24 which shows the distribution of $(B^2 - D^2)/\max(B^2, D^2)$. Near the equator the predominantly radial electric field is larger than the magnetic field and drives the electric current across the poloidal magnetic field lines. Both the radial component (the middle panel of Fig. 9.24) of magnetic field and its azimuthal component exhibit a break in the equatorial plane on the scale of the current sheet. The most important result is shown in the left panel of Fig. 9.24: all magnetic field lines penetrating the ergosphere are forced into rotation in the same sense as the black hole irrespective of whether they eventually cross the event horizon or not. Along these field lines there are outgoing fluxes of both energy and angular momentum. Indeed, in steady state force-free magnetospheres the angular momentum flux is proportional to H_ϕ and the energy flux is proportional to $\Omega_F H_\phi$. Notice that Ω and H_ϕ are the same in the Boyer–Lindquist as in the Kerr–Schild coordinates. As one can see in Fig. 9.24, both these quantities are nonvanishing along the field lines penetrating the ergosphere. Within the current sheet the electromagnetic energy and angular momentum are not conserved and, thus, the numerical results suggest that

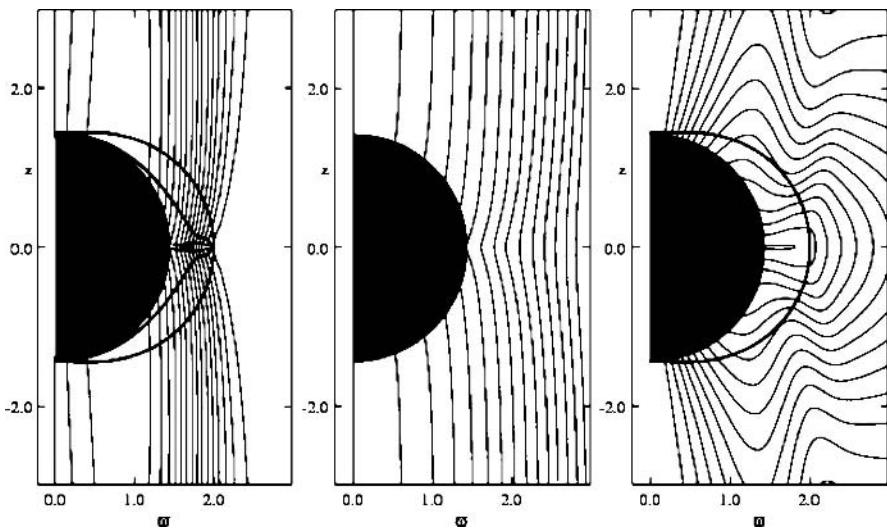


Fig. 9.24. Magnetospheric Wald problem. *Left panel:* The angular velocity of magnetic field lines. There are 15 contours equally spaced between 0 and 0.67. The angular velocity first gradually increased towards the axis but then reaches a maximum and goes slightly down. The thick lines show the ergosphere (the outer line) and the inner light surface (the inner line). *Middle panel:* The magnetic flux surfaces. *Right panel:* The distribution of $(B^2 - D^2)/\max(B^2, D^2)$. There are 15 contour equally spaced between -0.12 and 1.0 . This quantity monotonically decreases towards the current sheet in the equatorial plane within the ergosphere. The thick line shows the ergosphere. Figure adapted from Komissarov [238]

it is the current sheet that supplies both the energy and the angular momentum for the force-free magnetosphere above and below the sheet.

There is much more to the electrodynamics of black holes than it is proposed in the membrane paradigm and the analogy between a magnetized rotating conducting sphere and the black hole horizon is at least incomplete. In fact, closer inspection of the causality arguments due to Punsly and Coroniti shows that it is the membrane paradigm which is most directly under attack. The Blandford-Znajek solution itself is involved mainly because it is widely considered as inseparably linked with the paradigm.

On the Nature of the BZ Mechanism

It is well known that a black hole with zero total electric and magnetic charge cannot have its own magnetic field and, ultimately, any magnetic field penetrating the hole's ergosphere has to be supported by external currents. This fact alone makes black holes very different from magnetized stars like pulsars. Moreover, a steady state axisymmetric vacuum electromagnetic field cannot be used to extract energy and angular momentum of a rotating black hole. Indeed, the steady-state vacuum equations,

$$\nabla \times \mathbf{H} = 0, \quad \nabla \times \mathbf{E} = 0 \quad (9.153)$$

ensure

$$H_\phi = E_\phi = 0 \quad (9.154)$$

for axisymmetric configurations. Then one finds that the poloidal components of the energy and the angular momentum flux vectors vanish as well. In order to extract energy and angular momentum, the electromagnetic field has to be modified by magnetospheric charges and currents. However, in order to drive such currents and, perhaps, even to create charged particles via pair cascade in the first place, the electric field should not be screened. That is at least one of the above conditions has to be broken in the vacuum solution. As it is well known, the vacuum solution due to Wald has this property, but we need to know whether it is generic.

Thus, the vacuum field created within a black hole ergosphere by distant sources must have unscreened electric fields capable of driving electric currents, provided the charged particles are injected there somehow. The most popular mechanism involves an e^\pm pair cascade in strong electric and radiation field. Let us consider an initial vacuum solution of the kind discussed above, that is a solution with $\mathbf{D} \cdot \mathbf{B} = 0$ and $\mathcal{Q}_F = 0$, and figure out what occurs when plasma is injected.

First of all, FIDO's electric field $\mathbf{E} = \mathbf{D}$ will drive the conductivity current across the magnetic field lines within the ergosphere. This will result in the electric charge separation and the drop of the electrostatic potential along the magnetic field lines entering the ergosphere (see Fig. 9.25). The induced parallel component of the electric field will drive the conductivity current along the magnetic field lines leading to the dynamo of H_ϕ according to the Ampère law $\nabla \times \mathbf{H} = 4\pi\mathbf{J}$. As one

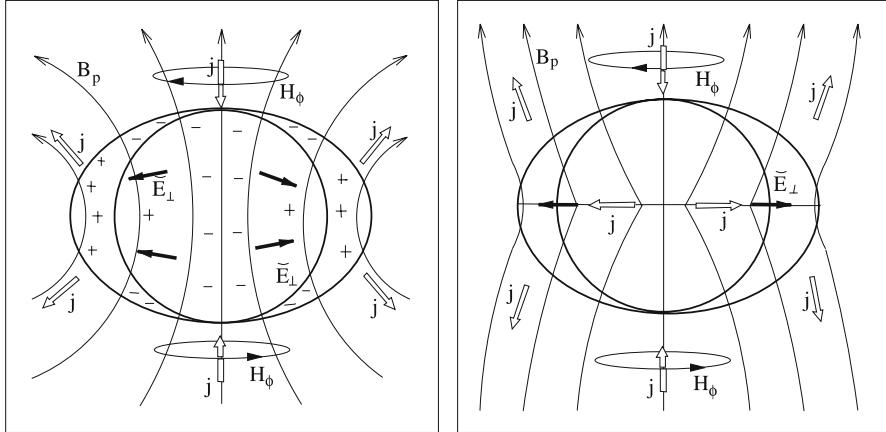


Fig. 9.25. Induction in the BH magnetosphere. *Left panel:* The initial currents driven by the gravitationally induced electric field within the ergosphere as it is being filled with pair plasma. *Right panel:* The electric current system of the steady-state Wald magnetosphere. Figure adapted from Komissarov [239]

can see in the left panel of Fig. 9.25, the sign of the generated H_ϕ is exactly the one which is required to slow down the black hole. However, we need to verify that this is not a temporary phenomenon and a nonvanishing H_ϕ , and, hence, a nonvanishing poloidal current, will be a property of the magnetosphere when it relaxes to a steady state.

Since the magnetosphere is driven to rotation in the ergosphere, the outgoing flux of angular momentum will inevitably result in rotation and, thus, extraction of the black hole energy as well. The equilibrium value of Ω_F depends of the details of the interaction. In the simulations [238], the black hole was surrounded by massless plasma and the equilibrium value of Ω_F was determined by the rate of deposition of energy and angular momentum in the surrounding space by means of propagating waves.

Even after reaching a force-free equilibrium, the magnetospheric electric field does not become completely screened. Within the ergospheric current sheet, E_\perp always remains slightly stronger than the magnetic field and keeps driving the cross field conductivity currents, thus, sustaining the potential drop along the magnetic field lines entering the current sheet. In the force-free region itself, it is the small residual component of E_\parallel that drives the poloidal conductivity currents, just as it occurs in the wires connected to the Faraday disk.

It is also quite clear why the magnetic field lines remaining outside of the ergosphere along their entire length do not rotate. Although the vacuum Wald solution has an unscreened E_\parallel component of the electric field which can trigger pair production and drive poloidal currents for a while, the injected charges eventually redistribute along the field lines and screen E_\parallel . Then, because E_\perp is too weak, even in the equatorial plane, to drive the cross field current, the poloidal currents die out completely.

The force free conditions are met even when $H_\phi = 0$. This conclusion is altered when the field lines outside the ergosphere are linked up with the surrounding disk. Then a smooth transition between the ergospheric rotation and the disk rotation is expected (as outlined in Fig. 9.25).

Where is the Unipolar Inductor of a Kerr Black Hole?

In the case of the Faraday disk, which is a classical example of a unipolar inductor, electrons are forced to participate in the disk rotation via collisions with other disk particles. This provides the electromotive force, $q\mathbf{v} \times \mathbf{B}$, which results in electric charge separation and, hence, the voltage drop between the disk rim and its center. When the disk is used as battery in a closed electric circuit the electromotive force continues to push electrons against the electric force and across the magnetic field lines. This is essential for sustaining the potential drop and for providing the current closure. Spinning magnetized cosmic objects like stars or accretion disks generate an electric field in the very much same way. Such a field is often described as *rotationally induced*.

Although, the membrane paradigm invites to treat black holes in a similar fashion, in reality they are rather different. Both in the Blandford–Znajek solution and in the solution to the magnetospheric Wald problem, there is no massive conducting rotating object and, thus, no usual electromotive force driving electric current over its surface. In spite of this curious feature the potential drop across the magnetic field lines still exists and the electric currents still flow.

The answer to this paradox seems to reside in another peculiar property of rotating black holes – contrary to our everyday experience, the electric charge separation is not the only way of creating stationary electric field in their vicinity. The vacuum solution found by Wald shows that such an electric field can be induced *gravitationally*. Unless a battery is continuously recharged via the action of some electromotive force, it cannot drive a stationary electric current. In other words, the initial transient currents can redistribute electric charge in the black hole magnetosphere in such a way that the electric field becomes completely screened and can no longer drive the electric currents. However, as we have shown above, such a final state is not possible within the black hole ergosphere. Black holes do not allow stationary solutions with screened electric field and vanishing poloidal electric currents.

Thus, the electrodynamics of rotating black holes is very different from electrodynamics of usual magnetic rotators and their batteries operate on other principles than the classical unipolar inductor. The key role played by the ergosphere in black hole electrodynamics allows us to call it the driving force of the Blandford–Znajek mechanism.

The Role of the Horizon

Since Blandford and Znajek employed the Boyer–Lindquist coordinate system, the horizon appears as a singular boundary of their spatial domain and the Znajek

condition appears as a boundary condition. In fact, this perception of the Znajek's condition is fully accommodated in the Membrane paradigm, where this condition is used to endow the event horizon with the properties of rotating conducting surfaces, reintroducing the missing unipolar inductor in a somewhat ghostly form.

Znajek's horizon condition is not a boundary condition after all [72]. Indeed, Znajek derived the horizon condition from a very natural requirement. Namely, the electromagnetic field at the event horizon had to be nonsingular when measured by a local free-falling observer. However, in the limit of force-free electrodynamics, the event horizon coincides with the fast critical surface of the ingoing wind. This immediately follows from the fact that the fast wave of force-free electrodynamics propagates with the speed of light. All this suggests that Znajek's condition is in fact a regularity condition at the fast critical point of the ingoing wind, obtained by Znajek in a somewhat unconventional way.

There is an important difference between a boundary condition and a regularity condition. Boundary conditions are set on the boundaries of computational domain and select particular solutions to both steady-state and time-dependent problems. Regularity conditions apply only to steady-state problems when one is looking for solutions passing smoothly through critical points, at which the steady-state equations change their type.

In the Kerr–Schild coordinates, where there is no coordinate singularity at the horizon, the critical nature of the Znajek condition becomes very clear. Let us consider a steady state force-free solution smoothly passing through the event horizon. In such a solution the angular velocity of magnetic field, Ω_F , and H_ϕ are constant along the magnetic field lines. We can use these constants in order to find the relationship between B^ϕ and B^r along a given magnetic field line

$$\mathbf{H} = \alpha \mathbf{B} + \frac{1}{\alpha} [\boldsymbol{\omega}(\boldsymbol{\beta} \cdot \mathbf{B}) - \mathbf{B}(\boldsymbol{\beta} \cdot \boldsymbol{\omega} + \beta^2) + \boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{B})]. \quad (9.155)$$

From this equation we derive, after rather involved calculations, the following result

$$B^\phi = \frac{\alpha H_\phi - B^r \sin^2 \theta (2r\Omega - a)}{\Delta \sin^2 \theta}. \quad (9.156)$$

The denominator of the right-hand side of this equation vanishes at the event horizon, where $\Delta = 0$. For B^ϕ to remain finite, the numerator has to vanish as well and we obtain

$$H_\phi = \frac{\sin^2 \theta}{\alpha_+} (2r_+ \Omega - a) B^r = \frac{(2r_+ \Omega - a) \sin \theta}{r_+^2 + a^2 \cos^2 \theta} A_{\phi,\theta}, \quad (9.157)$$

which is exactly the Znajek condition (notice that H_ϕ , A_ϕ , and Ω are the same in the Kerr–Schild coordinates as in the Boyer–Lindquist coordinates). Thus, Znajek's boundary condition is indeed a regularity condition at the fast critical point of the ingoing wind. In MHD approximation, in the case where the particle inertia is not negligibly small, the fast surface is located outside of the event horizon and

Znajek's horizon condition is no longer needed to determine the steady-state wind solution [100].

Once we have seen that (1) the event horizon does not play the role of a unipolar inductor, (2) Znajek's horizon condition is just the usual regularity condition, and (3) the key role in the electrodynamic mechanism is played by the black hole ergosphere, this dispute about the BZ mechanism has to be somewhat redirected. What we need to verify is that the ergosphere is causally connected with the outgoing wind. In fact, the inner Alfvén surface is always located inside the ergosphere, and the numerical simulations are fully consistent with this result. All these arguments allow us to conclude that there is no causality clash associated the electrodynamic mechanism in general and with the Blandford–Znajek solution in particular.

9.6 Magnetic Spin-Down of Rotating Black Holes

One consequence of this gravitomagnetic coupling is the fact that a rapidly rotating black hole generates a battery voltage drop near the horizon which then drives a huge current system through a magnetosphere. For this reason, we now integrate the induction law along a closed curve connecting a load region with the ergosphere through the disk

$$\oint_C \alpha \mathbf{E}_p \cdot d\mathbf{s} = -\frac{d}{dt} \int_A \mathbf{B} \cdot d\mathbf{S} - \oint_C (\boldsymbol{\beta} \times \mathbf{B}) \cdot d\mathbf{s}. \quad (9.158)$$

The last term acts as an additional term to Faraday's induction law for moving conductors. In a stationary situation we find the EMF

$$EMF(C) \equiv \oint_C \alpha \mathbf{E} \cdot d\mathbf{s} = -\oint_C (\boldsymbol{\beta} \times \mathbf{B}) \cdot d\mathbf{s}. \quad (9.159)$$

The integral along the flux surface outside the disk gives no contribution, and far away the gravitomagnetic effect drops out rapidly. There remains therefore only the contribution near the horizon. The main driving force is the integral near the horizon

$$\begin{aligned} EMF &= -\int (\boldsymbol{\beta}_H \times \mathbf{B}_\perp) \cdot d\mathbf{s} \simeq \frac{1}{2\pi} \Omega_H \Delta\Psi_H \\ &\simeq 10^{20} \text{ volt} \left(\frac{a}{M} \right) \left(\frac{M_H}{10^9 M_\odot} \right) \left(\frac{B_H}{10 \text{ kG}} \right), \end{aligned} \quad (9.160)$$

where $\Delta\Psi_H$ denotes the magnetic flux that covers the ergosphere. This EMF which drives the current system near the horizon depends strongly on the rotational state Ω_H of the black hole. It vanishes for nonrotating holes.

The maximum power which can be extracted from the rotational energy of the hole depends quadratically on the mass [80]

$$P_{H,\max} \simeq 10^{45} \text{ erg s}^{-1} \left(\frac{a}{M} \right)^2 \left(\frac{B_H}{10^4 G} \right)^2 \left(\frac{M_H}{10^9 M_\odot} \right)^2. \quad (9.161)$$

Such a system drives a current I_H that is also given by the rotation of the field lines

$$I_H \simeq \frac{1}{2} (\Omega_H - \Omega^F) B_H r_H^2 \simeq 10^{18} \text{ amp}. \quad (9.162)$$

As in the case of Jupiter's and of stellar magnetospheres, this current flow leads to a braking down of the rotating hole (for $\Omega_H > \Omega^F$) on a typical time-scale given by

$$t_{\text{spin-down}} = \frac{J_H}{dJ_H/dt} \simeq 1 \text{ Gyr} \left(\frac{B_H}{10 kG} \right)^{-2} \left(\frac{M_H}{10^9 M_\odot} \right)^{-1}. \quad (9.163)$$

Since the maximum field strength B_H also scales as $B_H \propto 1/\sqrt{M_H}$ (Eddington field strength), the braking time is practically independent of the mass M_H ! This is an interesting result: when supermassive black holes are born as rapidly rotating objects at redshift $z \simeq 2-6$, they would still be moderately rotating nowadays.

Let us discuss another application of this magnetospheric coupling. The toroidal velocity \mathbf{v}_T with respect to ZAMOs is

$$\mathbf{v}_T = \frac{\Omega - \omega}{\alpha} \mathbf{e}_\phi \quad (9.164)$$

and, instead of using the toroidal magnetic field, it is more convenient to introduce the current stream function $I = \alpha \varpi B^\phi$. In terms of these quantities, the poloidal electric field for a stationary configuration can be written as

$$\alpha \mathbf{E}_p = \frac{\eta}{\bar{\omega}} \nabla I \times \mathbf{e}_\phi - \frac{I}{\bar{\omega}} \mathbf{v}_p \times \mathbf{e}_\phi - (\Omega - \omega) \nabla \Psi. \quad (9.165)$$

η is the plasma diffusivity. The poloidal electric field is the result of currents and poloidal as well as toroidal plasma motions.

Let us consider a situation where magnetic fields are advected in the disk and superposed with some external field. A dipolar field is advected inwards and closes perpendicular to the horizon. In this case, the last term in the expression for the poloidal field vanishes at the horizon, where $\Omega(r_+) = \Omega_H = \omega(r_+)$, however in such a way that $(\Omega - \omega)/\alpha$ remains finite at the horizon. This term only creates a θ -component in the electric field. The rapid radial inflow also only generates a component of this type. The only radial component of the electric field is produced by the first term due to magnetic diffusivity. For a diffusivity vanishing near the horizon, those component will also disappear near the horizon so that currents that are flowing mostly radial in the disk will be deflected away from the disk near the horizon. Outside the disk, η vanishes so that the last term is the dominant contribution to the electric field (which is always perpendicular to the flux surfaces).

Problems

9.1. Spin Evolution of Black Holes: Under accretion, the spin of a black hole evolves according to (see Sect. 8.5.1)

$$\frac{da}{dt} = \frac{\dot{M}_{\text{acc}}}{M_H} [j_0(a) - 2a e_0(a)] + 2a \frac{\mathcal{P}}{M_H c^2} - \frac{\Omega_H}{8} (1 - \Omega_F/\Omega_H) \frac{r_H c^2}{GM_H} \frac{B_H^2 r_H^3}{M_H c^2}. \quad (9.166)$$

Show that $j_0(a) - 2ae_0(a) \rightarrow 0$ for $a \rightarrow 1$, when the inner edge of the Keplerian disk is at the marginal stable orbit.

Integrate the evolution $a(t)$ by including the Blandford–Znajek process with $\Omega_F = \Omega_H/2$ and for a fixed ratio of magnetic energy to black hole energy, $B_H^2 r_H^3 / M_H c^2 = \text{const.}$

9.2. Merging of Two Black Holes: Describe the various phases in the merging of two black holes in the center of a galaxy. Assume that the more massive black hole sits in the center of the spheroid. Which process brings in the second black hole from the parsec-scale to the subparsec-scale?

Calculate the crash-time for the black hole binary due to gravitational wave losses. Describe the form of the gravitational waves expected from this process.

9.3. A Cluster of Black Holes in the Galactic Center: Estimate the number of stellar black holes formed in the Bulge of our Galaxy. Discuss the possible effects for the stars in the Bulge.

9.4. Primordial Black Holes: Calculate the Hawking temperature for solar mass type black holes. Estimate their evaporation time-scale.

9.5. Spectrum of Sgr A*: Compile from the literature an energy spectrum for the Galactic center (radio, submillimeter, IR, X-rays and gamma-rays). Give a possible interpretation for the various spectral branches. Estimate the corresponding luminosity and compare with the Bondi accretion rate from the Galactic center region.

9.6. Light Cylinder Surfaces: Solve the equation for the inner and outer light cylinder surface around a rotating Kerr black hole. Plot the functions $r_L = r_L(\theta)$.

9.7. Poynting Flux in Force-Free Magnetospheres: Derive from the equations for force-free magnetospheres expressions for the Poynting flux and angular momentum flux in the magnetosphere.

9.8. Stationary MHD and Disk Winds: Use the membrane formalism to derive the five constants of motion for plasma flows in axisymmetric magnetospheres: the field rotation $\Omega_F(\Psi)$, the plasma load $\eta(\Psi)$, the total energy $E(\Psi)$, the total angular momentum $L(\Psi)$, and the entropy of the plasma flow $s(\Psi)$.

Solve these expressions for the Lorentz factor, the specific angular momentum and the toroidal magnetic field of the plasma flow as a function of the Mach number $M^2 = 4\pi\mu\eta$.

Investigate the critical points in the plasma flow and compare this with force-free conditions.

Derive from the normalization of the four-velocity the wind equation as a function of the Mach number M .

10 Physics of Accretion Flows around Compact Objects

The dominant accretion process for compact objects involves disk accretion. Gas supplied by donor stars in binary systems as well as by the host galaxy in quasars possesses sufficient angular momentum to form a disk-like structure around the compact objects. Accretion onto black holes in quasars powers the most luminous sources in the Universe. Quasars have now been detected up to redshifts of 6.4. In nearby galaxies, there is however not much fuel available so that extremely weak sources result from this accretion. Similarly, quiescent compact objects in binary systems show completely different signatures in the spectra compared to objects at high accretion rates, such as Cygnus X-1.

The physical state of accretion disks around compact objects will not depend very much on the mass of the central object, one of the fundamental parameters is the accretion rate in units of the Eddington accretion rate. Objects in the Universe differ by at least 10 orders of magnitude in their accretion rate. For the innermost structure of the resulting accretion disk, the spin of the central object may play some crucial role. In the modern view, an accretion disk is a quasistationary solution of radiative magnetohydrodynamics, provided the initial configuration has sufficient gas, angular momentum and magnetic fields. A slightly magnetized toroidal configuration has turned out to be a suitable test-bed as initial configuration. It provides the gas supply, is stabilized by angular momentum and winds up magnetic fields by differential rotation.

One of the basic questions in the theory of accretion disks is the origin of the angular momentum transport in disks. Turbulence in the disks is the natural driver for transporting angular momentum to larger radii. This is the basic assumption already made by the pioneering papers on accretion disks by Shakura and Sunyaev [363]. In the last years, turbulence driven by magnetic instabilities (MRI) turned out to be a potent source of angular momentum transport (Balbus and Hawley [48, 49]). Although magnetohydrodynamic (MHD) disk turbulence has long been seen as a source for angular momentum transport, current confidence rests on two more recent developments. Balbus and Hawley [48] demonstrated that the introduction of a weak magnetic field renders accretion flows linearly unstable. Numerous subsequent computer simulations showed then in the 1990s that the instability indeed rapidly develops into sustained turbulence, which transports angular momentum outwards at a rate that is consistent with observational constraints derived in studies of accretion in mass transfer binary systems.

The existence of the MRI tells us that the ultimate model for accretion disks is magnetohydrodynamics (MHD) including some form of radiation transport. The algorithmic treatment of classical MHD converges slowly towards a matured state, but the inclusion of radiative processes is still a big challenge for the future. In this chapter, we give an overview over the nature of MRI and its realization in computer simulations. For this purpose, we give an introduction into the algorithms used to simulate MHD processes on the computer, both in Newtonian and general relativistic approach.

This chapter is not meant to be a fundamental introduction into classical accretion theory. The reader is expected to be familiar with the concepts of standard accretion disks (also called α disks) and advection-dominated accretion flows (ADAFs). These topics can be found in many textbooks. It is the intention of the author to give an overview about modern concepts of accretion theory and observations, which are relevant for the future trends in the physics of accretion flows.

10.1 Angular Momentum Transport

In the absence of magnetic fields, a differentially rotating disk with angular velocity $\Omega(R)$ is linearly stable to axisymmetric perturbations according to the Rayleigh criterion if

$$\frac{d}{dR} (R^2 \Omega) > 0, \quad (10.1)$$

i.e. if the specific angular momentum $j(R) = R^2 \Omega$ of the flow is an increasing function of radius. For geometrically thin disks with Keplerian angular velocity $\Omega_K(R) = \sqrt{GM/R^3}$ the specific angular momentum $j(R) \propto \sqrt{R}$, and the disk is hydrodynamically stable. Numerical simulations seem to support this conclusion. Real disks are quite often geometrically thick with velocity profiles different from Keplerian ones. But they too are probably stable by the Rayleigh criterion.

This situation changes drastically, if the disk contains a magnetic field. Analytic studies have shown that a weak magnetic field destabilizes astrophysical disks, provided that

$$\frac{d\Omega^2}{d \ln R} < 0. \quad (10.2)$$

This condition is almost always satisfied in real disks. This **magnetorotational instability** (MRI) exists regardless of the initial magnetic field configuration [3, 48, 49, 407]. However, the growth rates will somewhat depend on the magnetic field configuration.

The main ideas behind the MRI can be derived from a simplified treatment already given by Balbus and Hawley [49]. We consider a fluid element orbiting in a disk with central gravitational potential $\Phi(r)$. We assume that pressure forces

are negligible, which is certainly not true for real disks. In cylindrical coordinates (R, z, ϕ) , the equations of motion read as

$$\ddot{R} - R\dot{\phi}^2 = -\partial_R \Phi + f_R \quad (10.3)$$

$$R\ddot{\phi} + 2\dot{R}\dot{\phi} = f_\phi. \quad (10.4)$$

f_R and f_ϕ are forces which will be specified in the following. We concentrate on a small patch of the disk at radius R_0 that is corotating with the overall orbital angular velocity $\Omega(R)$. Then we can define a local Cartesian coordinate system (x, y) with

$$R = R_0 + x, \quad \phi = \Omega t + y/R_0. \quad (10.5)$$

In linear order we obtain from the above equations of motion

$$\ddot{x} - 2\Omega\dot{y} = -x \frac{d\Omega^2}{d \ln R} + f_x \quad (10.6)$$

$$\ddot{y} + 2\Omega\dot{x} = f_y. \quad (10.7)$$

These equations describe the epicyclic motion of a pressureless fluid element perturbed from equilibrium.

If a disk contains a weak magnetic field, perturbations to the fluid in the plane of the disk will be opposed by magnetic tension forces generated by the bending of the field lines. We consider perturbations of the form $\exp i(\omega t - kx)$ with a corresponding magnetic tension force $\mathbf{f} = -(kv_A)^2 \mathbf{s}$, where \mathbf{s} is the displacement vector and $v_A = \sqrt{B_z^2/4\pi\rho}$ is the Alfvén speed. Using this expression for the forces f_x and f_y , and assuming a time-dependence $\exp(i\omega t)$, the above two equations become a coupled system

$$-\omega^2 x - 2i\omega\Omega y = -x \frac{d\Omega^2}{d \ln R} - (kv_A)^2 x \quad (10.8)$$

$$-\omega^2 x + 2i\omega\Omega y = -(kv_A)^2 y. \quad (10.9)$$

Combining these two relations yields a **dispersion relation** that is quadratic in ω^2

$$\begin{aligned} \omega^4 - \omega^2 & \left[\frac{d\Omega^2}{d \ln R} + 4\Omega^2 + 2(kv_A)^2 \right] \\ & + (kv_A)^2 \left[\frac{d\Omega^2}{d \ln R} + (kv_A)^2 \right] = 0. \end{aligned} \quad (10.10)$$

As usual, the system is unstable if $\omega^2 < 0$, which occurs when the third term in the dispersion relation is negative. Instability occurs therefore if

$$\frac{d\Omega^2}{d \ln R} + (kv_A)^2 < 0. \quad (10.11)$$

For sufficiently weak fields, $v_A \rightarrow 0$, or for perturbations of long enough wavelengths, $k \rightarrow 0$, we obtain the criterion mentioned in equation (10.2). In a real disk,

the longest wavelength perturbation in the vertical direction will be of the order of the scale-height H of the disk. There will be therefore instability, provided that the magnetic field is weaker than some threshold value B_{\max} .

The physical reason for this instability is fairly simple. Let us consider the above situation and consider the effect of perturbing a weak vertical field threading an otherwise uniform disk. If the field remains frozen into the plasma, field lines connecting adjacent annuli in the disk will be sheared by the differential rotation into a trailing spiral pattern. Provided the field is weak enough, magnetic tension will not keep the field lines back to the vertical. The magnetic tension acts to reduce the angular momentum of the inner fluid element and boost that to the outer one, providing angular momentum transport in the outward direction that is required to drive the accretion process.

In a Keplerian disk, the growth rate ω_{\max} can be as large as $3\Omega/4$, meaning that the growth of the magnetic field occurs on a dynamical time-scale. Analytic studies are however unable to follow the nonlinear evolution of the instability. One of the main question is here, what happens when the MRI reaches a kind of saturation rate and turbulence is excited throughout the entire flow? The important information is then contained in the stress tensor

$$W_{R\phi} = \varrho \langle v_R(v_\phi - R\Omega) - B_R B_\phi / (4\pi) \rangle, \quad (10.12)$$

which includes both the Reynolds and magnetic (Maxwell) stresses. The angle brackets denote a density weighted average over height. $W_{R\phi}$ can be measured in numerical simulations and it can be used to estimate the α parameter introduced by Shakura and Sunyaev [363] in the form of

$$W_{R\phi} = \alpha \varrho c_S^2, \quad (10.13)$$

where c_S is the local sound speed and α is a dimensionless parameter that measures the efficiency of angular momentum transport in disks.

Numerical Simulations of MRI

Numerical simulations of the nonlinear development of the MRI have now been performed in both local and global geometries. Local simulations follow the evolution of the MRI in a small patch of the disk with periodic boundary conditions in ϕ and R . Global simulations [38, 199] have the advantage of being able to study the large-scale magnetic fields and are essential for the investigation of of geometrically thick flows, where $H \simeq R$. All these simulations have one problem in common: The physical separation between the largest relevant scale (that of the disk) and the dissipative scale (either viscous or resistive) is always too large to be resolved numerically. In addition, in most simulations purely numerical effects are responsible for the dissipation. In all simulations it has been found that the MRI leads on a time-scale of just a few orbital periods to a sustained MHD turbulence in which angular momentum is transported outwards. The magnetic field energy stays in general smaller than the thermal energy in the disk. The Maxwell stress dominates by a large factor over the Reynolds stress. Values for α have been found in the range of 0.01 to 0.1.

10.2 Magnetohydrodynamics for Accretion Disks

The equations which have to be implemented on the computer are given either in conservative or nonconservative form. In the spirit of MRI, no viscosity is needed for the description of accreting plasmas. **An accretion disk is a quasistationary solution of radiative MHD for a given initial configuration with sufficient gas, angular momentum and magnetic fields.**

10.2.1 Equations of Magnetohydrodynamics

Due to the complexity, accretion onto compact objects can only be handled numerically. An understanding of the essential features of MHD is therefore crucial for compact objects. Many codes can be found on the market which solve these equations.

Advection Form for Radiative MHD

Accretion disks are global solutions of the equations for radiative MHD, no viscosity is needed to drive the angular momentum transport. According to the philosophy of MRI, angular momentum transport is self-consistently given by the MHD system. This system has been formulated, for example by Stone, Mihalas and Norman in 1992 [378], with $D/dt = \partial_t + \mathbf{v} \cdot \nabla$ as the advective derivative,

$$\frac{D\varrho}{dt} + \varrho(\nabla \cdot \mathbf{v}) = 0 \quad (10.14)$$

$$\begin{aligned} \varrho \frac{D\mathbf{v}}{dt} &= -\nabla P - \varrho \nabla \Phi \\ &\quad - \frac{1}{8\pi} \nabla \mathbf{B}^2 + \frac{1}{4\pi} (\mathbf{B} \cdot \nabla) \mathbf{B} + \frac{\chi\varrho}{c} \mathbf{F} \end{aligned} \quad (10.15)$$

$$\begin{aligned} \varrho \frac{D(e/\varrho)}{dt} &= -P(\nabla \cdot \mathbf{v}) - \kappa\varrho(4\pi B - cE) \\ &\quad + \frac{\eta}{16\pi^2} |\nabla \times \mathbf{B}|^2 + \nabla \cdot (\kappa \nabla T) + \sigma : \nabla \mathbf{v} \end{aligned} \quad (10.16)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left(\mathbf{v} \times \mathbf{B} - \frac{\eta}{4\pi} \nabla \times \mathbf{B} \right) \quad (10.17)$$

$$\varrho \frac{D(E/\varrho)}{dt} = -\nabla \mathbf{v} : \Pi_R - \nabla \cdot \mathbf{F} + \kappa\varrho(4\pi B - cE) \quad (10.18)$$

$$\mathbf{F} = -\frac{c\lambda_{\text{FLD}}}{\chi\varrho} \nabla E \quad (10.19)$$

$$\Phi(r) = -\frac{GM}{r - R_S} \quad (10.20)$$

$$\begin{aligned} \sigma_{ik} &= -l_T^2 \varrho \min(0, \nabla \cdot \mathbf{v}) \times \left(\nabla \mathbf{v}_{ik} - \frac{1}{3} \nabla \cdot \mathbf{v} \right) \\ &\quad + l_A \varrho (\delta x_i \times \min(0, \nabla v_{ik})^2) \delta_{ik}. \end{aligned} \quad (10.21)$$

Here ϱ , \mathbf{v} , e , and P are the gas density, velocity, internal energy density, and pressure, respectively, and \mathbf{B} is the magnetic induction. The total opacity χ is the sum of electron scattering $\sigma = 0.4 \text{ cm}^2 \text{ g}^{-1}$ and the free-free absorption opacity $\kappa = 10^{52} \varrho^{9/2} e^{-7/2} \text{ cm}^2 \text{ g}^{-1}$. E is the energy density of the radiation field, B the energy density in the Planck distribution, Π_R the pressure and \mathbf{F} the flux in the radiation field. Gas cools by emitting photons at a rate proportional to the black-body value $B = \sigma_B T^4 / \pi$, where σ_B is the Stefan–Boltzmann constant, $T = \mu \varrho / (\mathcal{R} \varrho)$ the gas temperature, $\mu = 0.6$ and \mathcal{R} the gas constant. The flux limiter λ_{FLD} is equal to 1/3 in optically thick regions (Levermore and Pomraning [252]). Causality is preserved in regions where the radiation energy density varies over optical depth less than unity by reducing the limiter to zero. An ideal gas equation is assumed, $P = (\Gamma - 1)e$ with $\Gamma = 5/3$.

Gravity is treated in the pseudo-Newtonian limit by means of the Paczynski–Wiita approach. This guarantees the same structure for the ISCO as in Schwarzschild, but neglects the correct boundary conditions for the angular velocity at the horizon (see the discussion in Sect. 8.5). The viscous stress tensor σ_{ik} is composed of artificial viscosity, which contains the tensor artificial viscosity and the artificial viscosity due to Neumann and Richtmyer (1950). l_T and l_A are the corresponding shock smearing length-scales. Advective schemes of this type are implemented in the ZEUS3D code [378] and NIRVANA2.0 [431]. These codes use a staggered grid to advance the solution of the induction equation. This method guarantees that $\text{div}(\mathbf{B}) = 0$ is satisfied with machine accuracy. More recent public-domain codes, such as FLASH [163] or PLUTO [291, 292], prefer to use the MHD equations in conservative form.

Conservative Form of Ideal MHD

In a real physical system, we should have the conservation of momentum, energy and mass. But the above set of equations cannot guarantee strict numerical conservation of momentum and energy, though it can assure the conservation of mass. At the same time, this scheme also has numerical difficulties with convective derivatives. The practical application of this set of equations also shows that nonconservative equations lead to numerical difficulties with strong shocks and to errors in the Rankine–Hugoniot conditions and shock speed [152].

The equations of ideal MHD can be formulated in true conservation form

$$\partial_t \varrho + \nabla \cdot (\varrho \mathbf{v}) = 0 \quad (10.22)$$

$$\partial_t (\varrho \mathbf{v}) + \nabla \cdot \left[\varrho \mathbf{v} \otimes \mathbf{v} + \left(P + \frac{1}{8\pi} \mathbf{B}^2 \right) \mathbf{I} - \frac{1}{4\pi} \mathbf{B} \otimes \mathbf{B} \right] = 0 \quad (10.23)$$

$$\partial_t \mathbf{B} - \nabla \cdot [\mathbf{v} \otimes \mathbf{B} - \mathbf{B} \otimes \mathbf{v}] = 0 \quad (10.24)$$

$$\partial_t (\varrho e) + \nabla \cdot \left[\left(\varrho e + P + \frac{\mathbf{B}^2}{8\pi} \right) \mathbf{v} - \frac{1}{4\pi} (\mathbf{v} \cdot \mathbf{B}) \mathbf{B} \right] = 0, \quad (10.25)$$

subject to the constraint $\nabla \cdot \mathbf{B} = 0$. The total energy density E is given by

$$E = \varrho e = \frac{P}{\Gamma - 1} + \frac{1}{2} \varrho \mathbf{v}^2 + \frac{1}{8\pi} \mathbf{B}^2. \quad (10.26)$$

When magnetic fields vanish, the ideal MHD equations reduce to the conservative form of the Euler equations for hydrodynamics (see Sect. 3.1).

This set of equations allows strict numerical conservation of mass, momentum and energy. But practical application of it shows that, in the region of low plasma β , where $\beta = 8\pi P_g / B^2$, numerical difficulties will be met. Sometimes pressure becomes negative, because P is the difference of large numbers (here we can also see that normalization has to be combined with other techniques to avoid the difference of large numbers, though normalization itself can prevent most such cases). The model of ideal MHD is therefore a **fully conservative system**, except for the induction equation which requires some regulation for maintaining numerically $\nabla \cdot \mathbf{B} = 0$ [131]. With the definition of the eight-dimensional state vector

$$\mathbf{U} = (\varrho, \varrho \mathbf{v}, \mathbf{B}, E)^T, \quad (10.27)$$

these equations can be written for Cartesian coordinates in true conservative form without source terms

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = 0.$$

(10.28)

The flux vector \mathbf{F} in general depends nonlinearly on the state variable \mathbf{U} , $\mathbf{F} = \mathbf{F}(\mathbf{U})$. We can then apply Godunov methods by integrating over a spacetime volume $\Delta t \times \Delta x^1 \times \Delta x^2 \times \Delta x^3$ to advance the state vector

$$\mathbf{U}(t + \Delta t) = \mathbf{U}(t) - \frac{\Delta t}{\Delta V} \sum_{\text{faces j}} \mathcal{F}^j \Delta S_j + \delta(g) \Delta t. \quad (10.29)$$

$\delta(g)$ is a source term due to non-Cartesian coordinates (see below). The summation is taken over all spatial faces of the cell, \mathbf{U} is the vector of conserved variables averaged over a cell volume, and \mathcal{F}^j are the fluxes averaged over the cell faces. In a Godunov scheme, the fluxes are found by solving the Riemann problem at the cell interfaces [396].

Equations in Curved Coordinates

The above equations can easily be formulated for Cartesian systems

$$(x^1, x^2, x^3) = (x, y, z), \quad (h_1, h_2, h_3) = (1, 1, 1), \quad (10.30)$$

spherical polar coordinates with

$$(x^1, x^2, x^3) = (r, \theta, \phi), \quad (h_1, h_2, h_3) = (1, r, r \sin \theta) \quad (10.31)$$

and cylindrical coordinates

$$(x^1, x^2, x^3) = (R, z, \phi), \quad (h_1, h_2, h_3) = (1, 1, R), \quad (10.32)$$

where h_i are metric scale factors. The conservative form of the MHD equations is slightly different for different geometries. They can still be written in semiconservative form, which now has a source term on the right-hand side,

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{1}{h_1 h_2 h_3} \left(\frac{\partial}{\partial x^1} [h_2 h_3 \mathbf{F}] + \frac{\partial}{\partial x^2} [h_1 h_3 \mathbf{G}] + \frac{\partial}{\partial x^3} [h_1 h_2 \mathbf{H}] \right) = \mathbf{S}. \quad (10.33)$$

The state vector is now

$$\mathbf{U} = (\varrho, \varrho v_1, \varrho v_2, \varrho v_3, B_1, B_2, B_3, E)^T \quad (10.34)$$

and the flux functions are explicitly given¹

$$\mathbf{F} = \begin{pmatrix} \varrho v_1 \\ \varrho v_1^2 - B_1^2 + P^* \\ \varrho v_1 v_2 - B_1 B_2 \\ \varrho v_1 v_3 - B_1 B_3 \\ 0 \\ \mathcal{E}_3 \\ -\mathcal{E}_2 \\ (E + P^*)v_1 - B_1(\mathbf{v} \cdot \mathbf{B}) \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} \varrho v_2 \\ \varrho v_2 v_1 - B_2 B_1 \\ \varrho v_2^2 - B_2^2 + P^* \\ \varrho v_2 v_3 - B_2 B_3 \\ -\mathcal{E}_3 \\ 0 \\ \mathcal{E}_1 \\ (E + P^*)v_2 - B_2(\mathbf{v} \cdot \mathbf{B}) \end{pmatrix} \quad (10.35)$$

and

$$\mathbf{H} = \begin{pmatrix} \varrho v_3 \\ \varrho v_3 v_1 - B_3 B_1 \\ \varrho v_3 v_2 - B_3 B_2 \\ \varrho v_3^2 - B_3^2 + P^* \\ \mathcal{E}_2 \\ -\mathcal{E}_1 \\ 0 \\ (E + P^*)v_1 - B_3(\mathbf{v} \cdot \mathbf{B}) \end{pmatrix}. \quad (10.36)$$

The source terms \mathbf{S} are quite complicated and given by derivatives of the metric functions (Christoffel symbols)

¹ For the following B_i is replaced by $B_i/\sqrt{4\pi}$.

$$\mathbf{S} = \begin{pmatrix} \frac{B_1 B_2 - \varrho V_1 v_2}{h_1 h_2} \frac{\partial h_1}{\partial x^2} + \frac{B_1 B_3 - \varrho V_1 v_3}{h_1 h_3} \frac{\partial h_1}{\partial x^3} \\ + \frac{\varrho V_2^2 - B_2^2}{h_1 h_2} \frac{\partial h_2}{\partial x^1} + \frac{\varrho V_3^2 - B_3^2}{h_1 h_3} \frac{\partial h_3}{\partial x^1} + \frac{P^*}{h_1 h_2 h_3} \frac{\partial [h_2 h_3]}{\partial x^1} \\ + \frac{B_2 B_1 - \varrho V_2 v_1}{h_1 h_2} \frac{\partial h_2}{\partial x^1} + \frac{B_2 B_3 - \varrho V_2 v_3}{h_2 h_3} \frac{\partial h_2}{\partial x^3} \\ + \frac{\varrho V_1^2 - B_1^2}{h_1 h_2} \frac{\partial h_1}{\partial x^2} + \frac{\varrho V_2^2 - B_2^2}{h_2 h_3} \frac{\partial h_3}{\partial x^2} + \frac{P^*}{h_1 h_2 h_3} \frac{\partial [h_1 h_3]}{\partial x^2} \\ + \frac{B_3 B_2 - \varrho V_3 v_2}{h_1 h_2} \frac{\partial h_3}{\partial x^2} + \frac{B_1 B_3 - \varrho V_1 v_3}{h_1 h_3} \frac{\partial h_3}{\partial x^1} \\ + \frac{\varrho V_1^2 - B_1^2}{h_1 h_2} \frac{\partial h_1}{\partial x^2} + \frac{\varrho V_2^2 - B_2^2}{h_2 h_3} \frac{\partial h_2}{\partial x^3} + \frac{P^*}{h_1 h_2 h_3} \frac{\partial [h_1 h_2]}{\partial x^3} \\ \frac{\mathcal{E}_2}{h_1 h_3} \frac{\partial h_1}{\partial x^3} - \frac{\mathcal{E}_3}{h_1 h_2} \frac{\partial h_1}{\partial x^2} \\ \frac{\mathcal{E}_3}{h_1 h_2} \frac{\partial h_2}{\partial x^1} - \frac{\mathcal{E}_1}{h_3 h_2} \frac{\partial h_2}{\partial x^3} \\ \frac{\mathcal{E}_1}{h_2 h_3} \frac{\partial h_3}{\partial x^2} - \frac{\mathcal{E}_2}{h_1 h_3} \frac{\partial h_3}{\partial x^1} \\ 0 \end{pmatrix}. \quad (10.37)$$

P^* is the total pressure

$$P^* = P + \frac{\mathbf{B} \cdot \mathbf{B}}{2}, \quad (10.38)$$

and \mathcal{E} the electromotive forces, $\mathcal{E} = \mathbf{v} \times \mathbf{B}$. The total energy E is given as previously by equation (10.26). This system consists of eight conservation equations, but only seven fluxes are nontrivial. The fluxes of B_i in the i -direction are always zero.

Eigenvalues and Eigenvectors

We also can write the conservation equations in primitive variables

$$\mathbf{P} = (\varrho, v_1, v_2, v_3, B_1, B_2, B_3, e)^T \quad (10.39)$$

in the following form, say in the x -direction,

$$\mathbf{A}_t \frac{\partial \mathbf{P}}{\partial t} + \mathbf{A}_x \frac{\partial \mathbf{P}}{\partial x} = 0,$$

(10.40)

where the matrices \mathbf{A}_i are the Jacobian matrices

$$\mathbf{A}_x(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}}, \quad \mathbf{A}_y(\mathbf{U}) = \frac{\partial \mathbf{G}}{\partial \mathbf{U}}, \quad \mathbf{A}_z(\mathbf{U}) = \frac{\partial \mathbf{H}}{\partial \mathbf{U}}, \quad \mathbf{A}_t = \frac{\partial \mathbf{U}}{\partial \mathbf{P}}. \quad (10.41)$$

This represents a hyperbolic system with the characteristic polynomial

$$|\mathbf{A}_x - \lambda \mathbf{A}_t| = 0. \quad (10.42)$$

\mathbf{A}_y and \mathbf{A}_z have similar structures as \mathbf{A}_x . The eigenvalues and eigenvectors have been studied by many authors [90, 346, 351]. There are two sets of eigenvectors for the eigensystem of \mathbf{A}_x . One is related to the one-dimensional system and the other is based on the modification proposed by Powell [331]. The fifth component is ignored

and the rest is identical to the one-dimensional flux where B_1 is a constant. The eigenvalues for the 7×7 system are

$$\lambda_{1,7} = v_1 \pm c_f, \quad \lambda_{2,6} = v_1 \pm c_A, \quad \lambda_{3,5} = v_1 \pm c_s, \quad \lambda_4 = v_1, \quad (10.43)$$

where

$$c_A = \sqrt{B_1^2/4\pi\varrho} \quad (10.44)$$

is the Alfvén wave speed, and

$$c_f = \sqrt{\frac{1}{2} \left(a^2 + \frac{B^2}{4\pi\varrho} + \sqrt{(a^2 + B^2/4\pi\varrho)^2 - 4a^2 c_A^2} \right)} \quad (10.45)$$

$$c_s = \sqrt{\frac{1}{2} \left(a^2 + \frac{B^2}{4\pi\varrho} - \sqrt{(a^2 + B^2/4\pi\varrho)^2 - 4a^2 c_A^2} \right)} \quad (10.46)$$

are the speeds of the fast and slow magnetosonic waves, and a is the acoustic wave speed defined as

$$a = \sqrt{\Gamma P/\varrho}. \quad (10.47)$$

In the 8×8 eigensystem of Powell, the eigenvalues are

$$\lambda_{1,8} = v_1 \pm c_f, \quad \lambda_{2,7} = v_1 \pm c_A, \quad \lambda_{3,6} = v_1 \pm c_s, \quad \lambda_{4,5} = v_1. \quad (10.48)$$

The eigensystem is discussed in Powell [331].

For a dimensional splitting integration scheme, the CFL conditions along each row is given by

$$\text{CFL} = \Delta t \times \min \left[\frac{|v_1| + c_f}{h_1 \Delta x^1}, \frac{|v_2| + c_f}{h_2 \Delta x^2}, \frac{|v_3| + c_f}{h_3 \Delta x^3} \right], \quad (10.49)$$

where the minimum is over all the cells.

Flux Formula

The Godunov method and its higher order extensions require the solution of the Riemann problem. In practical computations, this is solved billions of times, making the Riemann problem solution process the most demanding one in a numerical method. Numerical fluxes across the cell interfaces are computed by using various flux formula. One of the simplest implementation is the HLL flux formula (Harten et al. [194]) which is based on the maximum left- and right-propagating wave speeds

$$\mathbf{F}_{\text{HLL}} = \frac{\psi_+ \mathbf{F}(\mathbf{U}_L) - \psi_- \mathbf{F}(\mathbf{U}_R) + \psi_+ \psi_- (\mathbf{U}_R - \mathbf{U}_L)}{\psi_+ - \psi_-} \quad (10.50)$$

with

$$\psi_+ = \max(0, \lambda_{f,L}^+, \lambda_{f,R}^+), \quad \psi_- = \min(0, \lambda_{f,L}^-, \lambda_{f,R}^-). \quad (10.51)$$

The HLL solver has strong diffusion on the rarefaction waves and contact surfaces. Various improvements have been proposed.

These Riemann solvers are proposed for 1D problems and first-order method. To achieve higher order accuracy in space, a second-order or higher order reconstruction is required. The simplest case is a piecewise linear profile in the cell

$$U_i(x) = U_i + \frac{x - x_i}{\Delta x} \bar{\Delta u}_i, \quad (10.52)$$

where $\bar{\Delta u}_i$ is a limited slope for the cell. Typically, a minmod-slope limiter is used.

10.2.2 Time and Space Discretization

In the real world, space and time are continuous, but in an MHD simulation, we have to use discrete space and time to describe the system. The simplest way to discretize space is to divide the whole simulation region into many regions of the same size. For the time discretization, the simplest way is to choose the same time-step, Δt , for all the evolution. Though the uniform discretization of space and time is simple, we cannot hope that it can optimize the simulation efficiency and reliability. In order to have a better solution of this problem, we need finer grids in some regions, while at the same time, coarser grids are enough in other regions. A gridding scheme depending on the real physical component value, instead of being predefined, is called adaptive (AMR). The kind of step adjusting that can be done by program itself is called self-adaptive. The same thing can also be performed in time domain.

A package of FORTRAN 90 routines, called PARAMESH [264], has been developed to provide an application developer with an easy route to extend an existing serial code that uses a logically Cartesian structured mesh into a parallel code with AMR. PARAMESH builds a hierarchy of subgrids to cover the computational domain, with spatial resolution varying to satisfy the demands of the application (Fig. 10.1). These subgrid blocks form the nodes of a tree data structure (quad-tree in 2D or oct-tree in 3D). Each grid block has a logically Cartesian mesh, and the index ranges are the same for every block. Thus, in 2D, if we begin with a 10×20 grid on one block covering the entire domain, the first refinement step would produce four child blocks, each with its own 10×20 mesh, but now with mesh spacing $1/2$ that of its parent. Any or all of these children can themselves be refined in the same manner. This process continues until the domain is covered with a quilt-like pattern of blocks with the desired spatial resolution everywhere. A similar package called Chombo [120] provides a set of tools for implementing finite difference methods for the solution of partial differential equations on block-structured adaptively refined rectangular grids. Both elliptic and time-dependent modules are included. Support for parallel platforms and standardized self-describing file formats are included.

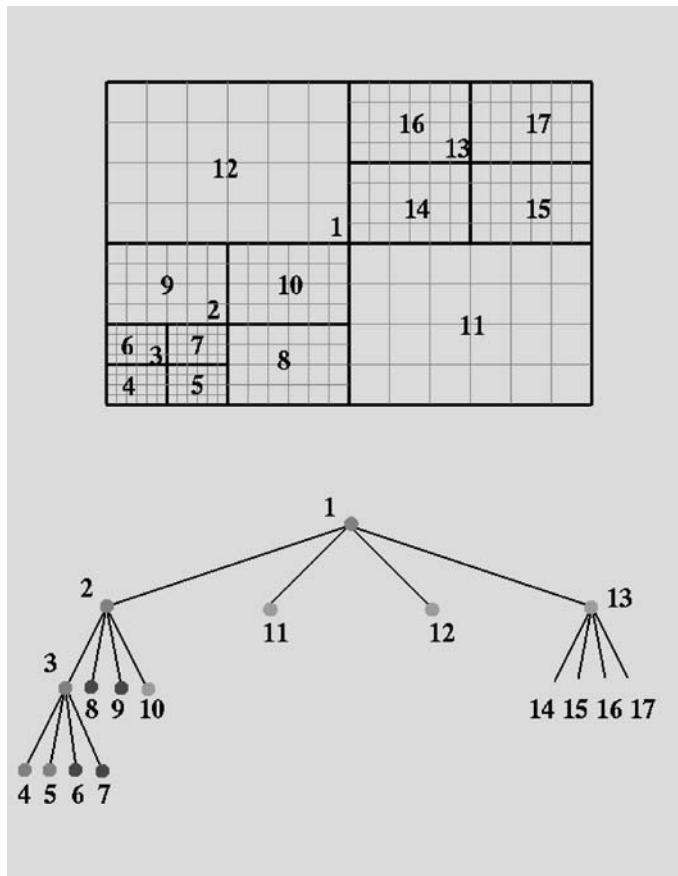


Fig. 10.1. A 2D 6×4 grid is created on each block. The *numbers* assigned to each block designate the blocks location in the quad-tree below. The *colors* assigned to the nodes of the tree indicate one possible distribution of the blocks during a four-processor calculation

Chombo provides a distributed infrastructure for parallel calculations over block-structured, adaptively refined grids.

Error Handling and Divergence Cleaning

The errors of analytical results come from the approximation of the theoretical model and the improper presentation of the physical environments and procedures. When a simulation model comes from a theoretical model, it also has errors generated by the analytical results. In addition, a numerical method has its own errors during the simulation process due to computation accuracy and imperfect representation of those continuous processes, such as differentiations in time and space. Simple error analysis can be conducted on the basis of each of the simulation equations, as shown

in the error terms [131]. Further error analysis of the whole system becomes much more prohibitive because of the complexity of the MHD equations. The comparison of simulation results with analytical results for simple problems can give us some very useful ideas of the errors of the MHD equation system.

Diffusive errors come from the diffusive property of MHD equations. In some cases, there are some sharp boundaries between two parts of the simulation region. But because of the error of the numerical computation, such sharp boundaries can be smoothed. Generally, diffusive errors smooth the parameter variation in space and make the spatial configuration more flat. In some cases, the error in a numerical computation can feed back on the simulation system. If the computation scheme cannot erase this error, some wiggles may appear. Sometimes these wiggles can be greatly enhanced through the feedback mechanism to an extent that the real physical process can be totally hidden.

In numerical MHD simulations the divergence constraint on the magnetic field causes severe stability problems. Accumulating errors can lead to an unphysical situation and can result in a breakdown of the simulation. Various authors have developed an approach to the stabilization of numerical schemes which can be easily used as an extension of an existing solver. The method is based on a modified formulation of the MHD equations in which the divergence constraint is coupled to the system by introducing a further unknown function. The evolution of divergence errors is strongly dependent on the type of the equation chosen for this function. For the one-dimensional setting, these errors can be transported out of the computational domain by a wave equation or can be dissipated by a heat equation. In [131] a mixed formulation is suggested, by which the divergence errors are transported and dissipated at the same time. The resulting system is still hyperbolic and the density, momentum, magnetic induction, and the total energy density are still conserved. Numerical examples demonstrate that this method decreases the divergence errors by up to two orders of magnitude even compared with the often used source term stabilization approach by Powell and coworkers [331].

10.2.3 MRI Driven Turbulence in Disks

As an application of the previously discussed methods we consider the evolution of the MRI in strong differentially rotating disks.

Linear Stability Analysis

Let us consider an equilibrium disk of a differentially rotating fluid in which there is a uniform vertical component of magnetic field, as well as an azimuthal (toroidal) component which is too weak to exert significant stresses on the equilibrium flow. We also assume that the flow is incompressible, $\nabla \cdot \mathbf{V} = 0$. We now consider local perturbations (i.e. wavelengths much less than equilibrium wavelengths) with wavenumbers purely in the vertical direction. The incompressibility condition then implies $\delta v_z = 0$, i.e. the **fluid motions are entirely horizontal**. Similarly, the

condition $\nabla \cdot \mathbf{B} = 0$ implies that there is no perturbation in the vertical component, $\delta B_z = 0$.

The perturbed equations resulting from the induction equation are then

$$\frac{\partial \delta B_R}{\partial t} = B_z \frac{\partial \delta v_R}{\partial z} \quad (10.53)$$

$$\frac{\partial \delta B_\phi}{\partial t} = R \frac{d\Omega}{dR} \delta B_R + B_z \frac{\partial \delta v_\phi}{\partial z}. \quad (10.54)$$

These equations describe the dynamics of the perturbed field. Vertical gradients of radial and azimuthal velocity stretch the equilibrium vertical field out into radial and azimuthal field, respectively. In addition, the background differential rotation shears out the perturbed radial field into azimuthal field.

The perturbed fluid momentum equations are

$$\frac{\partial \delta v_R}{\partial t} - 2\Omega \delta v_\phi = \frac{B_z}{4\pi\varrho} \frac{\partial \delta B_R}{\partial z} \quad (10.55)$$

$$\frac{\partial \delta v_\phi}{\partial t} + \frac{\kappa^2}{2\Omega} \delta v_R = \frac{B_z}{4\pi\varrho} \frac{\partial \delta B_\phi}{\partial z} \quad (10.56)$$

$$\frac{\partial \delta P}{\partial z} + \frac{B_\phi}{4\pi\varrho} \frac{\partial \delta B_\phi}{\partial z} = 0. \quad (10.57)$$

κ is the epicycle frequency given by

$$\kappa^2 = R \frac{d\Omega^2}{dR} + 4\Omega^2 = \frac{1}{R^3} \frac{dJ^2}{dR}. \quad (10.58)$$

The radial and azimuthal momentum equations are like the spring equations we discussed in the previous section. Vertical gradients in the azimuthal field create vertical magnetic pressure gradients (the last equation) which, for an incompressible flow, are immediately balanced by vertical fluid pressure gradients. This equation becomes trivial, if the equilibrium azimuthal field B_ϕ vanishes.

Assuming a spacetime dependence of the perturbations, $(\delta v, \delta B) \propto \exp[i(kz - \omega t)]$, the first four equations can be solved to give the **dispersion relation**

$$\omega^4 - [2(kV_A)^2 + \kappa^2] \omega^2 + (kV_A)^2 \left((kV_A)^2 + R \frac{d\Omega^2}{dR} \right) = 0. \quad (10.59)$$

$V_A = B_z / \sqrt{4\pi\varrho}$ is the Alfvén speed corresponding to the vertical field². This immediately implies that one of the solutions of this quadratic equation for ω^2 will be negative (implying an unstable mode), if and only if

$$(kV_A)^2 + R \frac{d\Omega^2}{dR} < 0. \quad (10.60)$$

² The most general dispersion relation is discussed by Balbus and Hawley [49].

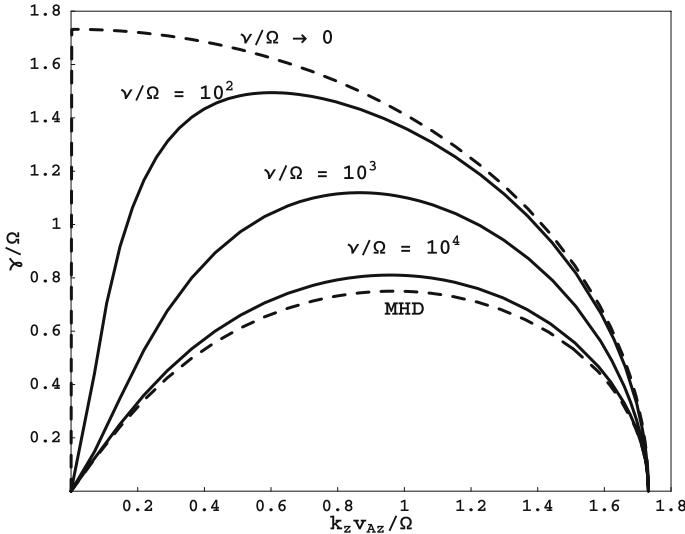


Fig. 10.2. MRI growth rate as a function of the vertical wavenumber, normalized to the Alfvén speed. The dispersion relation for ideal MHD is shown as a *dashed line*. *Solid lines* relate to the viscous case, parametrized by the ion-ion collision frequency ν in units of the rotational frequency Ω . Figure adapted from [214]

All accretion disks around black holes have angular velocities that decrease outwards. Hence, sufficiently small Alfvén speeds always catalyze an instability! It is then easy to show that the maximum instability growth rate, $\Gamma = i\omega$,

$$\Gamma_{\max} = \frac{1}{2} \left| \frac{R d\Omega}{dR} \right|, \quad (10.61)$$

occurs for wavenumbers (see Fig. 10.2)

$$(kV_A)^2_{\max} = \Omega^2 - \frac{\kappa^4}{16\Omega^2}. \quad (10.62)$$

For Keplerian disks we find $\Gamma_{\max} = 3\Omega_K/4$, $\kappa = \Omega_K$, and therefore

$$\left(\frac{kV_A}{\Omega_K} \right)_{\max} \simeq 1. \quad (10.63)$$

This implies that the instability grows on the Keplerian time-scale.

This magnetorotational instability (MRI) relies on weak magnetic fields. If the field is too strong at the considered wavelength, then magnetic tension will overcome the magnetic torquing and the flow will stabilize. The stronger the field, the longer the wavelength required for instability. This suggests that laminar flows are unstable, provided an unstable MRI wavelength $\simeq V_A/\Omega$ can fit inside the vertical thickness of the disk. This corresponds to initial magnetic field densities that are smaller than the thermal energy densities. This has been confirmed by global simulations.

Global MRI Simulations

Recent increases in supercomputer performance have significantly improved the ability to evolve the basic equations of accretion disk structure and evolution. These developments, along with continuing progress in understanding the most important physical processes that occur within accretion disks, suggest that predictive disk simulations are a realistic goal. Such disk simulations will be global, fully three-dimensional, and incorporate physical processes such as magnetohydrodynamics (MHD) and radiation transport. At present, we are still some ways from this goal; global simulations are still rather idealized in terms of disk structure, energetics, and dynamical range.

Most of the global disk simulations for black holes are nonrelativistic and assume as initial condition a polytropic gas torus in the gravitational field of a pseudo black hole. This means that the gravitational potential is given by the Paczynski–Wiita form [319]

$$\Phi(r) = -\frac{GM}{r - R_S}, \quad (10.64)$$

where R_S is the Schwarzschild radius. This initial condition is chosen to be a torus for various reasons. First, the mechanical equilibrium can easily be constructed, and, secondly, a torus is a kind of reservoir for angular momentum located at some large distance away from the black hole. This second fact is however also the crux for numerical simulations, because large dynamical ranges in radii are required. In

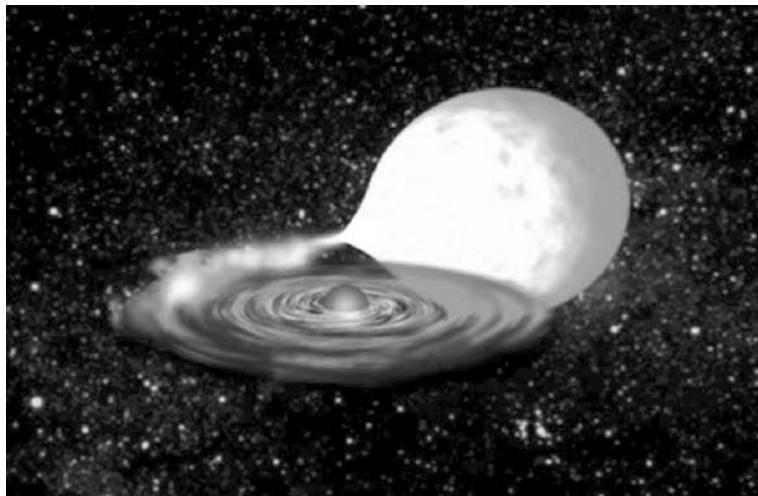


Fig. 10.3. Compact objects (white dwarfs, neutron stars and stellar black holes) as companion stars in binary systems accrete gas from a normal star. This gas carries over angular momentum, such that a turbulent accretion disk is formed. These disks are typically optically thick in the outer parts and optically thin near the compact object

a quasar host galaxy, this would require something like 100,000 dynamical radii. The Paczynski–Wiita potential is a useful approximation for nonrotating black holes, but it requires the wrong boundary conditions for differential rotation. In a real Schwarzschild background, the angular velocity $\Omega(r)$ should vanish along the horizon. This is a consequence of the no-rotation following from frame-dragging (see Sect. 8.4). In studies of accretion disks around neutron stars, one has to introduce a real boundary layer along the surface of the neutron star.

Much of this work has focused on geometrically thick accretion disks. With a pressure scale-height H comparable to the disk radius R , the thick disk, or accretion torus, is more easily resolved in a numerical simulation than disks for which $H/R \ll 1$. Matsumoto [275] followed the evolution of a thick torus embedded in an external vertical field, and found significant outflow collimated along the global vertical field lines. Hawley [199] considered tori containing toroidal fields and poloidal

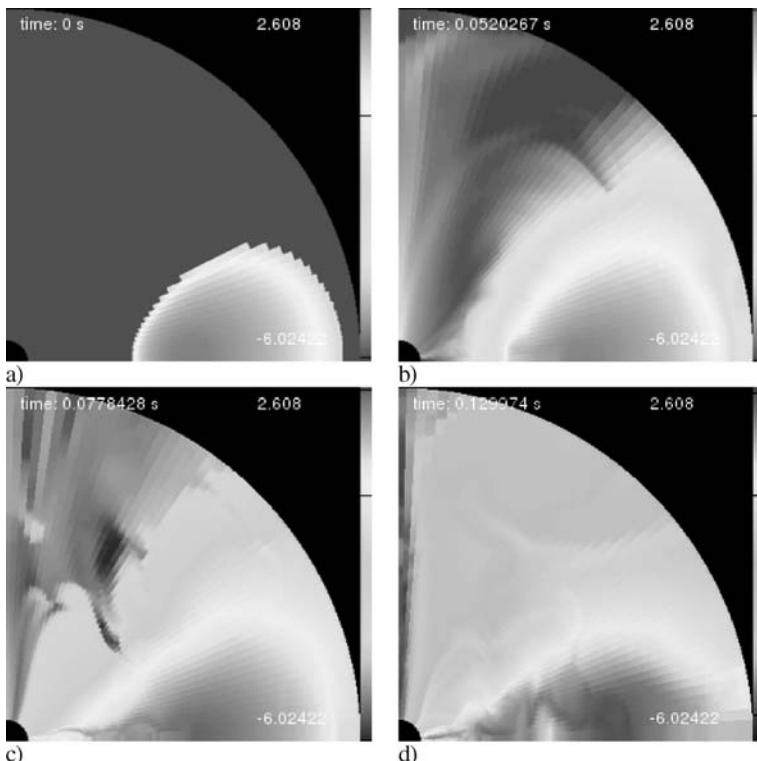


Fig. 10.4. Evolution of a torus including weak magnetic fields (plasma $\beta = 10^3$) for various time-steps. The black hole of 10 solar masses (*dark central region*) is treated in the pseudo-Newtonian approach. The density is represented in logarithmic scale. Shown is the original torus configuration (a) and the density distribution at one (b), 1.5 (c) and 2.5 (d) revolution periods at density maximum. White numbers (*right-hand side*) indicate the values of the density maximum and minimum. Simulation provided by S. Brinkmann [88], based on the NIRVANA code (ZAH, LSW Heidelberg)

field loops. In these studies the initial field was entirely contained within the disk and the resulting outflows were confined to the creation of a magnetized corona. A generic feature of all these thick disk simulations is the presence of large amplitude fluctuations in accretion rate, density, and other variables, in both space and time.

Thick accretion tori with initially non-Keplerian angular momentum distributions are highly unstable. MHD turbulence develops rapidly and is sustained by a self-consistent dynamo process within the disk (Fig. 10.4). The constant or near-constant specific angular momentum distribution of the initial torus rapidly evolves to one that is near Keplerian. In principle, the rotation profile $\Omega(R)$ is a free function, but in practice, almost all MRI simulations show that the Keplerian profile rapidly emerges (Fig. 10.5). The reason for this is the vigorous outward transport of angular momentum that is established by MRI. This spreads initially the torus radially, diluting the effects of pressure gradients. Balbus [51] has shown that for a polytropic equation of state, $P = K\varrho^\Gamma$, the equilibrium density profile satisfies

$$\frac{K\Gamma\varrho^{\Gamma-1}}{\Gamma-1} = h_\infty + \frac{GM}{r} \left(1 - \frac{\cos\beta}{\cos\lambda}\right), \quad (10.65)$$

where h_∞ is the enthalpy at infinity and λ is the colatitude. For an isothermal equation of state one obtains

$$\varrho = \varrho_\infty \exp\left[\frac{GM}{r} \left(1 - \frac{\cos\beta}{\cos\lambda}\right)\right]. \quad (10.66)$$

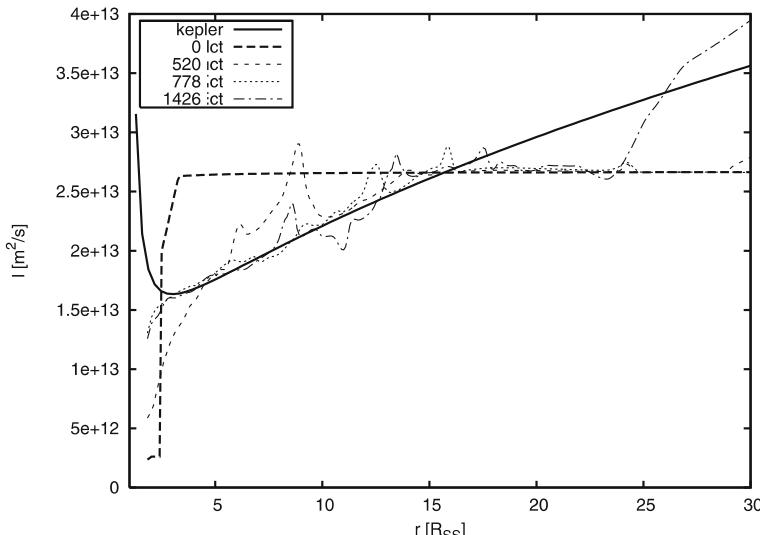


Fig. 10.5. Evolution of the angular momentum distribution in the accreting torus. Radii are given in units of Schwarzschild radii. The initial constant specific angular momentum (dashed line) quickly evolves towards a quasi-Keplerian distribution (solid line). Compare with Fig. 10.4. Figure provided by S. Brinkmann [88] (LSW Heidelberg)

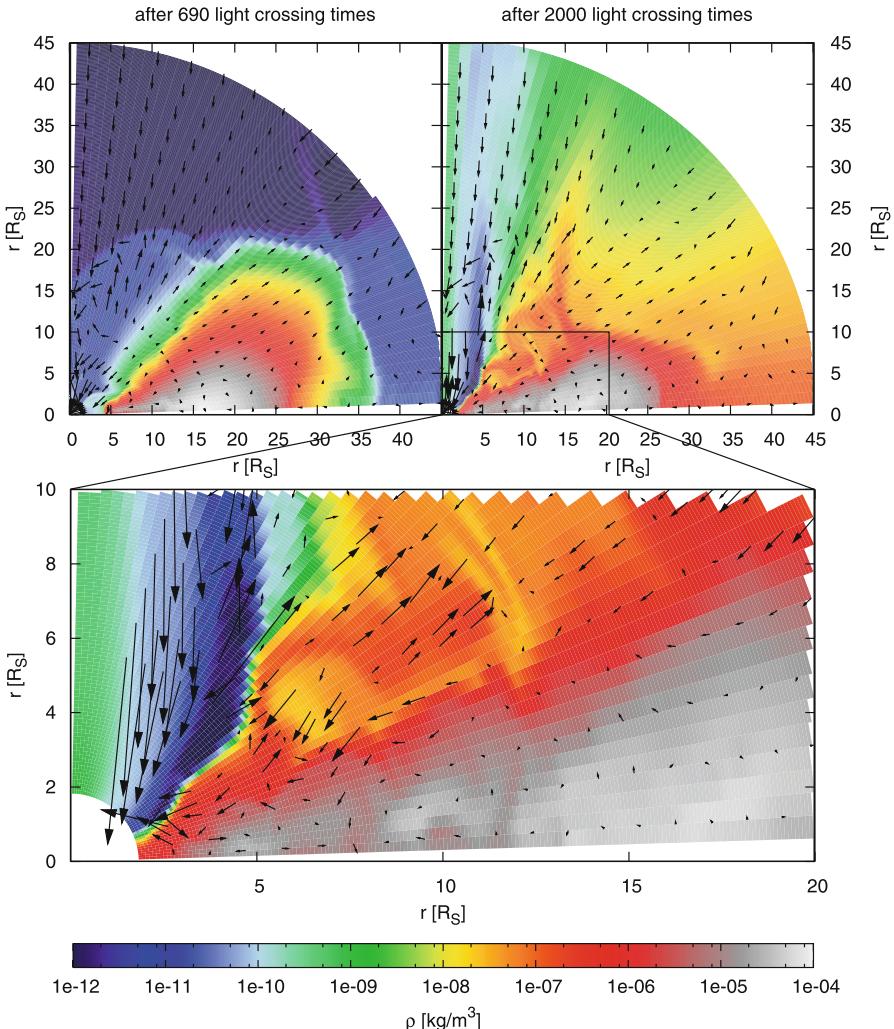


Fig. 10.6. The magnetorotational instability (MRI) in action. The images show the density distribution in the meridional plane of a nonrotating black hole together with velocity vectors. The initial torus was in hydrodynamical equilibrium with a plasma beta of 10^3 , but it evolves rapidly as the simulation proceeds. The radial field is sheared by the differential rotation, creating toroidal field. Turbulence develops within the disk with the onset of the magnetorotational instability (MRI). The resulting Maxwell stresses drive angular momentum transfer; the disk begins to accrete into the central hole and expands radially outward. One revolution at the torus maximum corresponds to 500 lct. *Top left:* initial phase; *top right:* evolution after four torus rotation periods; *lower panel:* zoom into the central part. Images based on simulations provided by S. Brinkmann [89] (ZAH, LSW Heidelberg)

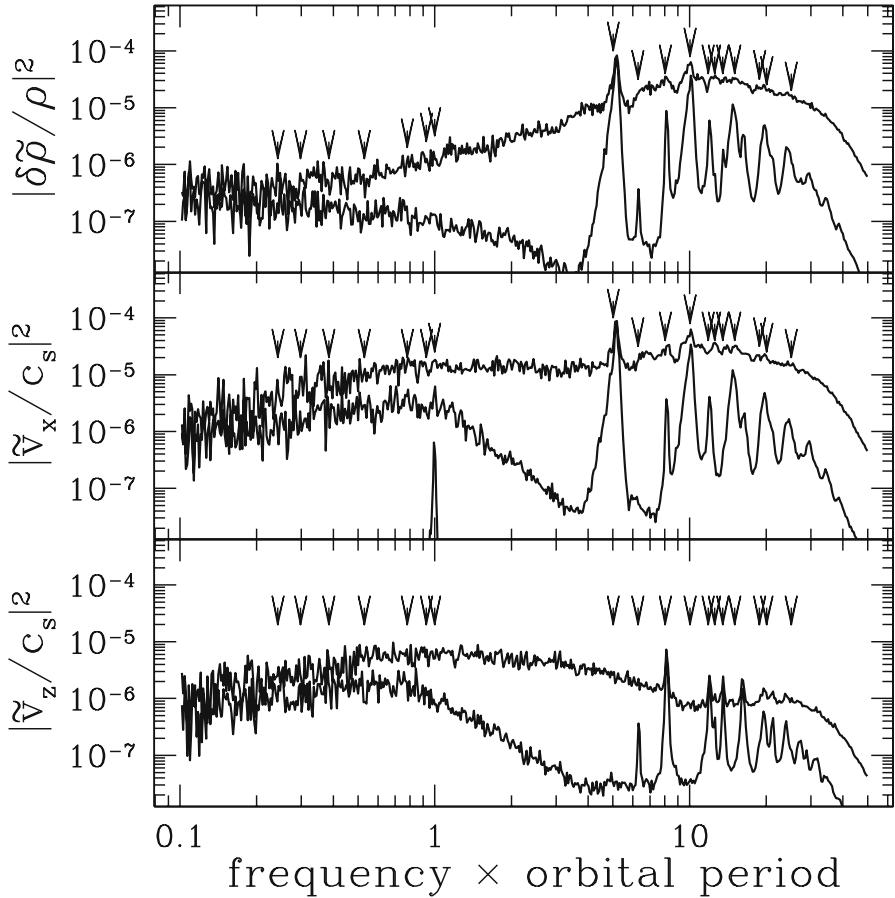


Fig. 10.7. Position space power spectrum for density (*upper*), radial velocity (*middle*), and vertical velocity fluctuations (*lower*). In each plot, the upper curve is for the nonaxisymmetric data and the lower curve is for the axisymmetric data. The sharp peak at the orbital frequency in the *middle* plot is the box averaged data. The frequency is in units of cycles per orbital period. Power has been rebinned into logarithmically spaced frequency bins, which effectively multiplies the power spectrum by one power of frequency. Axisymmetric hydrodynamic normal mode frequencies are shown as arrows. The epicyclic mode is at the orbital frequency (unity). Inertial modes and sound waves should lie below and above the epicyclic frequency, respectively. Figure adapted from Arras et al. [42]

The simulation shown in Fig. 10.4 assumes a toroidal configuration located at 16 Schwarzschild radii, having initially a constant specific angular momentum. The computational domain is in spherical coordinates (r, θ, ϕ) running from 2 to 30 Schwarzschild radii in r , $0 \leq \theta \leq \pi/2$. An outflow boundary condition is used at both the outer and inner radial boundaries, and periodic boundary conditions are used for ϕ .

Already after one revolution at the torus center, the MRI sets in, field energy is amplified, and soon the characteristic radial streaming structures (referred to as the channel solution) of the vertical field instability appear, much as they do in the local shearing box models. In the present simulation, these structures develop first at the inner part of the torus where the rotation frequency is the highest. The amplitude of the MRI becomes nonlinear by three orbits at the center of the grid, and filaments of strong magnetic field are carried inward and outward by fluid elements well out of Keplerian balance. These reach the outer part of the disk even before the local MRI in that region becomes fully nonlinear. Thus there are two immediate global effects not seen in local simulations: linear growth rates that vary strongly with radius ($\omega_{\text{MRI}} \simeq \Omega \simeq R^{-3/2}$), and extended radial motions of significantly non-Keplerian plasma.

Quasiperiodic Oscillations from Magnetorotational Turbulence

Quasiperiodic oscillations (QPOs) in the X-ray light-curves of accreting neutron star and black hole binaries have been widely interpreted as being due to standing wave modes in accretion disks. These disks are, however, thought to be highly turbulent due to the magnetorotational instability (MRI). Arras et al. [42] have recently studied wave excitation by MRI turbulence in the shearing box geometry. They demonstrate that axisymmetric sound waves and radial epicyclic motions driven by MRI turbulence give rise to narrow, distinct peaks in the temporal power spectrum (Fig. 10.7). Inertial waves, on the other hand, do not give rise to distinct peaks which rise significantly above the continuum noise spectrum set by MRI turbulence, even when the fluid motions are projected onto the eigenfunctions of the modes. This is a serious problem for QPO models based on inertial waves.

The most striking feature in Fig. 10.7 is the axisymmetric sound waves, which appear as sharply defined peaks. The power in these peaks approaches the broad continuum due to nonaxisymmetric MRI turbulence, and clearly rises above the continuum for at least two of the peaks.

These results demonstrate that QPOs can be found in simulations of MRI turbulence, at least in a shearing box geometry. The well-defined boundary conditions of this geometry help ensure the existence of modes, but it is not clear that such well-defined boundaries will exist in global accretion disks. Global MRI simulations could in principle determine self-consistently, whether wave cavities exist and contain trapped modes. This is one of the exciting prospects for future global MRI simulations.

10.2.4 Two-Temperature Plasmas and Radiation Pressure in Accretion Disks

Radiation is essential for the understanding of classical accretion disks. One of the basic assumptions was that all free energy locally dissipated by turbulence is radiated away. Very little has been done to include radiative losses in numerical simulations. Hujeirat and Rannacher [205] have pointed out that different time-scales have to be included when radiation is accounted for. This requires a modification in the

integration of the MHD equations, time-implicit codes are better suited to model the evolution of such systems.

Since ions and electrons are only weakly coupled in low density accretion disks, the equations governing radiative MHD for applications in accretion disks can become quite complicated [206]

$$\partial_t \varrho + \nabla \cdot (\varrho \mathbf{v}) = 0 \quad (10.67)$$

$$\begin{aligned} \partial_t (\varrho \mathbf{v}_p) + \nabla \cdot (\varrho \mathbf{v}_p \otimes \mathbf{v}_p) &= -\nabla P + \frac{1}{8\pi} \nabla \mathbf{B}^2 + \frac{1}{4\pi} (\mathbf{B} \cdot \nabla) \mathbf{B} \\ &\quad - \varrho \nabla \Phi + \lambda_{\text{FLD}} \nabla E + \boldsymbol{\Pi}_{\text{visc}} \end{aligned} \quad (10.68)$$

$$\partial_t (\varrho j) + \nabla \cdot (j \varrho \mathbf{v}) = \frac{1}{4\pi} \mathbf{B}_p \cdot \nabla (R B_\phi) + \boldsymbol{\Pi}_{\text{visc}}^\phi \quad (10.69)$$

$$\begin{aligned} \partial_t \mathcal{E}_i + \nabla \cdot (\mathcal{E}_i \mathbf{v}) &= -P_i(\nabla \cdot \mathbf{v}) + \frac{\eta}{16\pi^2} |\nabla \times \mathbf{B}|^2 \\ &\quad - \Lambda_{i-e} + \nabla \cdot (\kappa_i \nabla T) + \Phi_{\text{vis}} \end{aligned} \quad (10.70)$$

$$\begin{aligned} \partial_t \mathcal{E}_e + \nabla \cdot (\mathcal{E}_e \mathbf{v}) &= -P_e(\nabla \cdot \mathbf{v}) + \Lambda_{i-e} \\ &\quad - \Lambda_B - \Lambda_{\text{cyc}} - \Lambda_C \end{aligned} \quad (10.71) \quad (10.72)$$

$$\begin{aligned} \partial_t E + \nabla \cdot (E \mathbf{v}) &= \nabla \cdot [\lambda_{\text{FLD}} \nabla E] + \Lambda_B + \Lambda_C + \Lambda_{\text{cyc}} \\ &\quad - \Lambda_{i-e} + \nabla \cdot (\kappa_e \nabla T) \end{aligned} \quad (10.73)$$

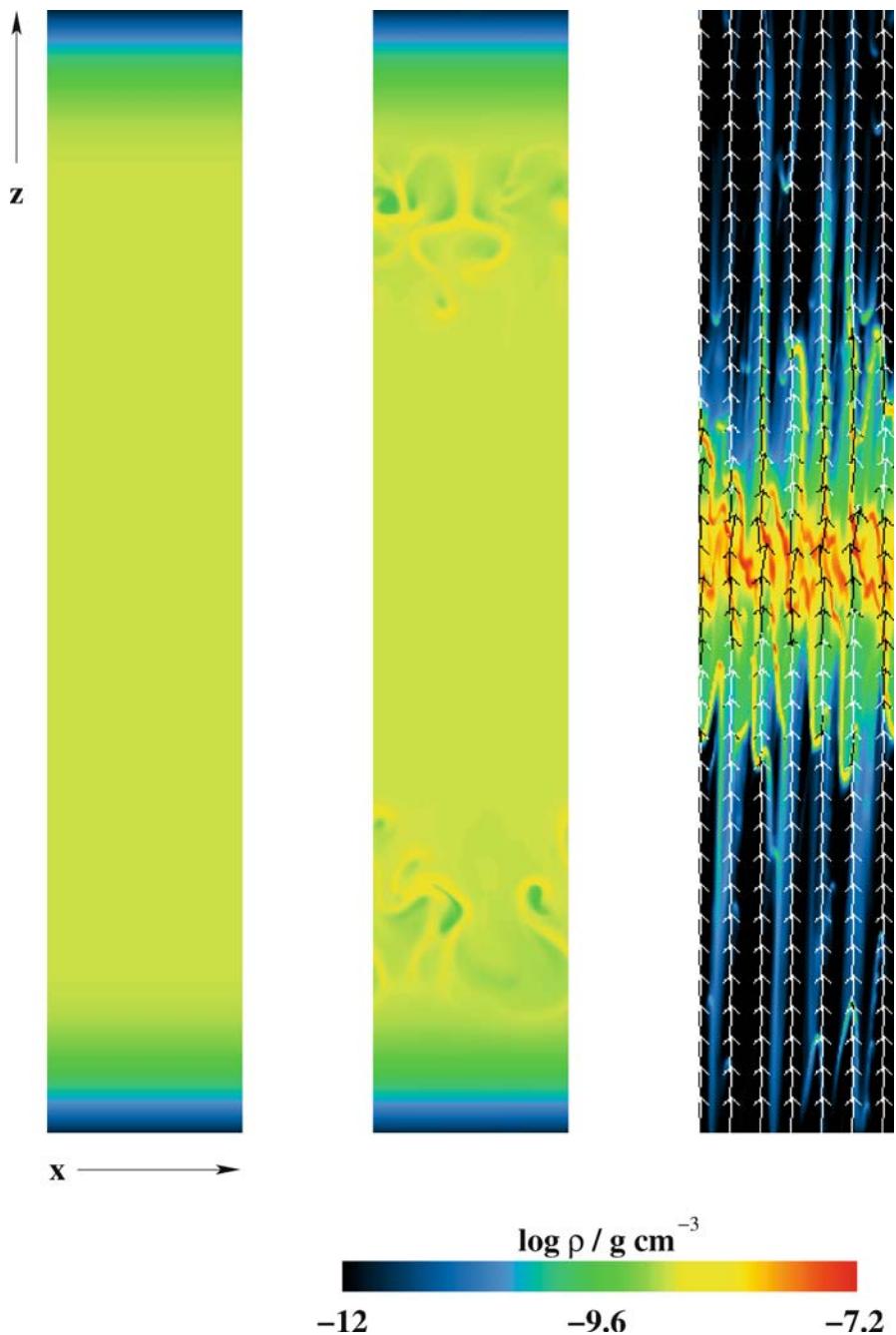
$$\partial_t \mathbf{B} = \nabla \times \left(\mathbf{v} \times \mathbf{B} - \frac{\eta}{4\pi} \nabla \times \mathbf{B} \right). \quad (10.74)$$

The subscripts i and e correspond to ions and electrons in the plasma, respectively. Radiation is treated as a first moment, i.e. by means of the radiative energy density E , and the radiative diffusion coefficient λ_{FLD} is a flux limiter which forces the radiative flux to adopt the correct form in an optically thin medium. We have also split the velocity vector into a poloidal component and the specific angular momentum $j = R^2 \Omega$, $\mathbf{v} = (\mathbf{v}_p, R\Omega)$. Λ_{i-e} , Λ_B , Λ_C and Λ_{cyc} correspond to electron heating by ions, Bremsstrahlung cooling, Compton and cyclotron cooling, respectively. These functions are given as follows [13] (in cgs units)

$$\Lambda_{i-e} = 5.95 \times 10^{-3} n_i n_e c k_B \frac{T_i - T_e}{T_e^{3/2}} / \mathcal{N} \quad (10.75)$$

$$\Lambda_B = 4a c \kappa_{\text{abs}} \varrho (T^4 - E) / \mathcal{N} \quad (10.76)$$

Fig. 10.8. The photon bubble instability. Snapshots of the gas density at 1.5 orbits in three calculations of a small patch of accretion disk, centered 20 Schwarzschild radii from a black hole of $10^8 M_\odot$. The domain extends 1.15 Shakura–Sunyaev scale-heights H above and below the midplane and the width is $0.4H$. At *left* are results from a 1D calculation in which the gas is cooled by vertical radiation diffusion and contracts slightly toward the midplane. At *center* in a 2D calculation, convection starts in the outer layers. The version shown at *right* differs only in including a magnetic field (*arrows*). The field is initially uniform, inclined 87° from horizontal, with pressure 10% of the midplane radiation pressure. Figure adapted from Turner et al. [403]



$$\Lambda_C = 4\sigma_T n_e c \frac{k_B}{m_e c^2} (T_e - T_{\text{rad}})/\mathcal{N}. \quad (10.77)$$

$\mathcal{N} \propto (\Gamma - 1)/\Gamma$ is a normalization quantity. The radiative temperature is defined as $T_{\text{rad}} = E^{1/4}$.

Magnetic Effects in Radiation-Dominated Accretion Disks

Black hole systems with luminosities above a few percent of the Eddington limit are thought to be powered by accretion through a geometrically thin disk supported by radiation pressure. Near the hole, the internal pressure greatly exceeds the gas pressure. The released gravitational energy escapes by diffusion of photons to the disk faces. This model is unstable to perturbations in mass flow and heating rates. The thermal instability is absent if additional cooling processes operate. In MRI models, magnetic fields linking material at different distances from the hole transfers angular momentum outward. The gas is heated by dissipation of magnetic fields and the turbulence through microscopic resistivity and viscosity. The strength of the magnetic fields is regulated by generation through MRI and losses through buoyancy and dissipation.

One of the basic questions is how radiation-dominated disks cool. Several dynamical instabilities have linear growth rates similar to or even faster than the orbital frequency [402]. This may lead to a vertical energy transport in turbulence driven by the MRI. The most important instabilities are convection, the Shaviv modes and photon bubbles (Fig. 10.8):

- The standard SS model is convectively unstable (Bisnovatyi-Kogan and Blinnikov [73]). Convection may be quenched if heating from the dissipation of magnetic fields is concentrated at low column depths (Turner [403]).
- Radiation-supported atmospheres through which photons diffuse in less than a sound-crossing time may be subject to a linear instability (Shaviv [365]). This might lead to the formation of low-density chimneys where the radiative flux is enhanced.
- Displacements of gas along magnetic field lines can be overstable, leading to growing and propagating density variations known as photon bubbles (Arons [41]; Gammie [164]). Growth is fastest at short wavelengths (Blaes and Socrates [75]). This photon bubble instability may lead to the development of trains of propagating shocks.

The effects of these instabilities can be simulated by using a flux-limited diffusion module for MHD, which involves conservation of mass, plasma momentum, radiation flux, plasma energy, and radiation energy, as well as the evolution of the magnetic fields, for a local shearing box, as formulated by Turner et al. [403]

$$\frac{D\varrho}{dt} + \varrho(\nabla \cdot \mathbf{v}) = 0 \quad (10.78)$$

$$\varrho \frac{D\mathbf{v}}{dt} = -\nabla P - \frac{1}{8\pi} \nabla \mathbf{B}^2 + \frac{1}{4\pi} (\mathbf{B} \cdot \nabla) \mathbf{B} + \frac{\chi\varrho}{c} \mathbf{F} - \varrho \Omega^2 \mathbf{z} \quad (10.79)$$

$$\varrho \frac{D(e/\varrho)}{dt} = -P(\nabla \cdot \mathbf{v}) - \kappa_P \varrho (4\pi B - cE) \quad (10.80)$$

$$\varrho \frac{D(E/\varrho)}{dt} = -\nabla \mathbf{v} : \Pi_R - \nabla \cdot \mathbf{F} + \kappa_P \varrho (4\pi B - cE) \quad (10.81)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) \quad (10.82)$$

$$\mathbf{F} = -\frac{c\lambda_{\text{FLD}}}{\chi\varrho} \nabla E. \quad (10.83)$$

The vertical component of gravity is included by an acceleration $g_z = \Omega^2 z$.

These equations have been simulated by Turner et al. [403] using the ZEUS MHD code with its FLD module developed by Turner and Stone [401]. In the absence of magnetic fields, convective instability grows at about the orbital frequency Ω and cools the disk much faster than photon diffusion (Fig. 10.8). Including magnetic fields, photon bubble modes with wavelengths shorter than the gas pressure scale-height grow faster than the orbital frequency in the disk surface layers. Disturbances reaching nonlinear amplitudes steepen into trains of shocks. These results indicate that photon bubbles may be important in cooling radiation-dominated accretion disks.

Ultraluminous X-ray sources (ULXs) with observed luminosities of $10^{39.5}$ to 10^{41} erg/s and spectra fitted by multitemperature black body radiation of temperatures 1–2 keV are thought to be black holes accreting mass from a late O type stellar companion at high rate. In order not to exceed the Eddington limit, the accreting black holes would have to be of intermediate mass, 50 – $100 M_\odot$, while the observed disk temperatures are too high for such black hole masses. Various suggestions have been made to resolve this problem. Radiation pressure supported disks in magnetic fields are unstable against a photon-bubble instability. In a nonlinear version, such radiation driven inhomogeneities can allow the escaping flux to exceed the Eddington limit and were suggested as resolution of the ULX problem. It is proposed that radiation pressure leads to a complete fragmentation of accretion disks in magnetic fields around black holes and that radiation escapes through the gaps between fragments at super-Eddington luminosity. Essential for the gap formation is the overlying disk corona into which gas of too low a density evaporates, similar to evaporation of hot gas into the interstellar medium.

10.3 States of Turbulent Accretion Disks

X-ray binaries thought to contain black holes show at least three spectral states [125] (see Fig. 1.5). The **low/hard state** (LHS) is characterized by a power law and an exponential cutoff at about 200 keV or a thermal Comptonization model with an electron temperature of about 70 keV. A quiescent state has been suggested to exist, but in reality, the X-ray and radio properties of low luminosity BH candidates

form a continuum down to the lowest observable luminosities. The spectrum of the **high/soft state** (HS) is dominated by a thermal component thought to arise in a standard thin disk (SSD). This state also exhibits a weak power-law tail without observable cutoff. The very **high state** (VHS) displays the same two components, but with the steep power-law dominating the total flux. Steady radio emission, which is likely coming from a jet, has been found in the low/hard state, and strong radio flares have been seen during the very high state [155]. No strong radio emission has ever been detected in the high/soft state. Radio emission from active nuclei is also correlated with the X-ray spectral properties, and therefore with the geometry of the inner accretion flow.

This suggests that there is a class of flows that radiate inefficiently near the surface of a black hole [21, 308] (Fig. 10.10). Analytic models of nonradiating accretion flows (generically called NRAsFs) have been based on the classical α prescription.

10.3.1 Turbulent Angular Momentum Transport in Accretion Disks

The purpose of this section is to elucidate the nature of turbulent transport in accretion disks. Weak-turbulence theory allows a connection to be made between a phenomenological transport modelling and the fundamental fluid equations. As pointed out by Balbus [51], this is equivalent to retaining the quadratic correlations in the turbulent fluctuations. Kato et al. [219] had already applied this formalism to black hole accretion.

Global Conservation Laws in Turbulent Flows

Accretion disks satisfy two global conservation laws: conservation of angular momentum and conservation of total energy. To investigate the behavior of a turbulent flow, we consider the time-steady form of angular momentum conservation

$$\nabla \cdot \left[\varrho R^2 \Omega \mathbf{v} - \frac{RB_\phi}{4\pi} \mathbf{B} \right] = 0 \quad (10.84)$$

and the energy conservation law

$$\nabla \cdot \left[\varrho \mathbf{v} \left(\frac{v^2}{2} + \Phi + \frac{\Gamma \Theta}{\Gamma - 1} \right) + \frac{1}{4\pi} \mathbf{B} \times (\mathbf{v} \times \mathbf{B}) \right] = -Q^- . \quad (10.85)$$

The second term in the energy equation represents the Poynting flux in the flow. Contributions from viscosity and thermal conduction have been neglected. We understand all equations as azimuthally averaged.

We next expand all quantities into mean plus fluctuating components. The mass flux is then

$$\langle \varrho \mathbf{v} \rangle = \langle \varrho \rangle \langle \mathbf{v} \rangle + \langle \delta \varrho \delta \mathbf{v} \rangle , \quad (10.86)$$

which satisfies mass conservation

$$\nabla \cdot (\varrho \mathbf{v}) = 0. \quad (10.87)$$

Then we may define a stress vector

$$\mathbf{W} = \langle \delta v_\phi \delta \mathbf{v} - \frac{B_\phi \mathbf{B}}{4\pi} \rangle. \quad (10.88)$$

In terms of these quantities, angular momentum conservation reads as

$$\boxed{\nabla \cdot [\langle \varrho R^2 \Omega \mathbf{v} \rangle + \langle \varrho \rangle R \mathbf{W}] = 0.} \quad (10.89)$$

Similarly, the energy equation, though second order in weak turbulence theory, assumes the form

$$\begin{aligned} \nabla \cdot & \left[\langle \varrho \mathbf{v} \rangle \left(\frac{R^2 \Omega}{2} + \Phi + \frac{\Gamma \langle \Theta \rangle}{\Gamma - 1} \right) \right. \\ & \left. + \langle \varrho \rangle R \Omega \mathbf{W} + \frac{\Gamma}{\Gamma - 1} \langle \varrho \rangle \langle \delta \Theta \delta \mathbf{v} \rangle \right] = -Q^- . \end{aligned} \quad (10.90)$$

Θ is the normalized temperature, $\Theta = k_B T / \mu$. The first group of terms corresponds to the Bernoulli flux, which in general need not be constant. The essential new term which is not included in standard thin disk theory is the correlated thermal energy flux $\langle \delta \Theta \delta \mathbf{v} \rangle$. In standard models, the dominant nonadvective energy flux comes from the rotational transport term \mathbf{W} . The ratio of thermal energy flux to rotational stress will be of order $c_S / R \Omega$.

When combining mass conservation with radial hydrostatic equilibrium

$$R \Omega^2 = \frac{1}{\varrho} \frac{\partial P}{\partial R} + \frac{\partial \Phi}{\partial R}, \quad (10.91)$$

the energy equation can be written as

$$\begin{aligned} \frac{\langle \varrho v_R \rangle}{2R^2} \frac{d(R^4 \Omega^2)}{dR} + \frac{\langle \Theta \rangle}{\Gamma - 1} \langle \varrho \mathbf{v} \rangle \cdot \nabla S \\ + \nabla \cdot \left(\langle \varrho \rangle R \Omega \mathbf{W} + \frac{\Gamma}{\Gamma - 1} \langle \varrho \rangle \langle \delta \Theta \delta \mathbf{v} \rangle \right) = -Q^- . \end{aligned} \quad (10.92)$$

Energy is transported through a turbulent disk by various processes

$$\boxed{- \left(W_{R\phi} \frac{d\Omega}{d \ln R} + Q_{\text{rad}} \right) = \nabla \cdot \mathbf{F}_{\text{th}} + \frac{\langle \Theta \rangle}{\Gamma - 1} \langle \varrho \mathbf{v} \rangle \cdot \nabla S,} \quad (10.93)$$

where Q_{rad} is the volumetric radiation loss rate, S is the mean entropy and \mathbf{F}_{th} is the thermal energy flux

$$\mathbf{F}_{\text{th}} = \frac{\Gamma \varrho}{\Gamma - 1} \langle \delta \mathbf{v} \delta \Theta \rangle . \quad (10.94)$$

The velocity–temperature correlation is responsible for convective transport. This energy equation states that the rate at which energy is locally extracted from differential rotation is equal to the rate at which it is carried away by waves, dissipated as heat, and radiated away.

The second term on the right-hand side of the energy equation (10.93) is the average rate at which entropy is generated by the mean mass flow. If the energy is not carried off by adiabatic processes, then it must be dissipated. In the absence of radiative losses, it must drive either an inflow or outflow. In nonradiative accretion flow simulations one sees strong outflows generated.

In standard accretion disk theory (SSD), both terms on the right-hand side of the energy equation (10.93) are ignored, only radiation losses are included

$$-W_{R\phi} \frac{d\Omega}{d \ln R} = Q_{\text{rad}} . \quad (10.95)$$

The energy equation is a crucial question in ZEUS-like codes, which base their energetics on internal, as opposed to a total energy equation in conservative schemes. The important question is how much free energy is locally radiated away and how much is transported elsewhere. This has to be carefully checked in future codes.

One-Dimensional Approaches

On dimensional reasons, one can approximate the functional forms for \mathbf{W} and for the thermal flux by means of

$$\langle \varrho \delta v_R \delta v_\phi - B_R B_\phi / 4\pi \rangle = \alpha_{\text{SS}} \langle P \rangle \quad (10.96)$$

$$\langle \varrho \delta v_R \delta \Theta \rangle = \alpha_T \langle P \rangle \langle c_S \rangle , \quad (10.97)$$

where α_{SS} (the Shakura–Sunyaev α parameter) and α_T are dimensionless constants. The dynamical stress must be positive to extract energy from the differential rotation, i.e. $\alpha_{\text{SS}} > 0$, whereas the thermal flux can have either sign.

10.3.2 Truncated Accretion and Standard Disk Models in 1D

Two families of models are currently considered to describe an accretion flow onto black holes and production of the observed X-ray radiation: (i) a standard cold accretion disk with a hot corona above it, and (ii) an outer truncated accretion disk with a hot semispherical inner flow (called ADAF). The general idea is that accretion in the outer parts is optically thick and then makes a sudden transition towards an optically thin state (Fig. 10.9). The radius, where this transition occurs, is called transition radius. This situation will be analyzed in the following in the 1D approximation.

advection-dominated accretion flow

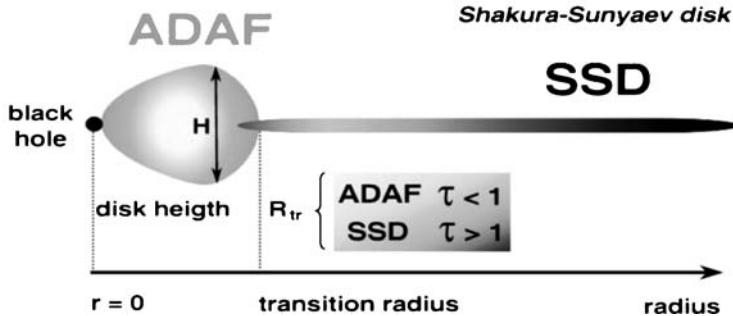


Fig. 10.9. Stationary thin disk solutions. Four types of solution branches are discussed in the literature: standard optically thick accretion disk for low accretion rates, slim disk solution for optically thick high accretion rates, the Shapiro–Eardley–Lightman disk for hot optically thin plasmas, and the advection-dominated ADAF solution. Figure adapted from A. Müller [304](LSW Heidelberg)

1D Disk Dynamics

For axisymmetric flows it is quite convenient to introduce vertically integrated variables, thereby reducing the 2D equations into a set of 1D equations. The basic variables are: the mass-density ϱ , the velocity field $\mathbf{V} = (V_R, V_\phi, 0)$ in cylindrical coordinates (R, ϕ, z) , the pressure p , and the internal energy e with the following expressions:

- surface mass density

$$\Sigma(t, R) = \int_{-H}^H \varrho(t, R, z) dz \simeq 2H \varrho(t, R, 0) \quad (10.98)$$

- integrated pressure

$$P(t, R) = \int_{-H}^H p(t, R, z) dz \simeq 2H p(t, R, 0) \quad (10.99)$$

- integrated stress for viscous processes

$$W_{R\phi}(t, R) = \int_{-H}^H \pi_{R\phi}(t, R, z) dz = \nu \Sigma R \frac{d\Omega}{dR} \quad (10.100)$$

- integrated dissipation

$$Q^+(t, R) = \int_{-H}^H \mathcal{D}(t, R, z) dz. \quad (10.101)$$

The system of conservation equations is now vertically integrated to result in the following system of 1D evolution equations

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\Sigma \mathbf{V}) = 0 \quad (10.102)$$

$$\begin{aligned} \frac{\partial(\Sigma V_R)}{\partial t} + \nabla \cdot (\Sigma V_R \mathbf{V}) &= -\frac{\partial P}{\partial R} + R \Sigma (\Omega^2 - \Omega_K^2) \\ &+ \frac{4}{3} \frac{\partial}{\partial R} \left(\frac{\nu \Sigma}{R} \frac{\partial(RV_R)}{\partial R} \right) + 2V_R \frac{\partial(\nu \Sigma)}{R \partial R} \end{aligned} \quad (10.103)$$

$$\frac{\partial(\Sigma j)}{\partial t} + \nabla \cdot (\Sigma j \mathbf{V}) = \frac{1}{R} \partial_R (R^2 W_{R\phi}) \quad (10.104)$$

$$\frac{\partial E}{\partial t} + \nabla \cdot (E \mathbf{V}) + P(\nabla \cdot \mathbf{V}) = Q_{\text{visc}} + Q_{\text{cond}} + Q_{\text{turb}} - Q_{\text{rad}}. \quad (10.105)$$

The first equation determines the surface mass density Σ , the second one the radial drift V_R , the third one the angular momentum distribution $j(t, R)$ and the last equation the internal energy E . Please note that in our case $\nabla \cdot (\Sigma \mathbf{V}) = (1/R) \partial_R (R \Sigma V_R)$. In this notation, the equations are written in conservative form.

Thermodynamics

In addition, we need an equation of state which includes in general radiation pressure

$$p = \frac{\varrho k_B T}{\mu m_u} + \frac{a}{3} T^4, \quad p_G = \beta p. \quad (10.106)$$

The equation for the internal energy is then

$$\varrho \mathcal{E} = \frac{\varrho k_B T}{\mu m_u (\Gamma - 1)} + a T^4 \quad (10.107)$$

with $\Gamma = C_p/C_V$ as the ratio of the specific heats. From this we obtain the equation of state for the integrated pressure

$$P = \frac{\Sigma k_B T}{\mu m_u} + \frac{1}{3} 2aHT^4. \quad (10.108)$$

We also can express the total variation of the pressure as

$$\begin{aligned} dp &= p_G (d \ln T + d \ln \varrho) + 4p_R d \ln T \\ &= \beta p (d \ln T + d \ln \varrho) + 4(1 - \beta)p d \ln T \\ &= (4 - 3\beta)p d \ln T + \beta p d \ln \varrho, \end{aligned} \quad (10.109)$$

and the total variation of the internal energy

$$\begin{aligned}
d\mathcal{E} &= \frac{p_G}{\varrho(\Gamma - 1)} d \ln T + 3 \frac{p_G}{\varrho} (4d \ln T - d \ln \varrho) \\
&= \frac{p}{\varrho} \left[\frac{\beta}{\Gamma - 1} d \ln T + 3(1 - \beta)(4d \ln T - d \ln \varrho) \right] \\
&= \frac{p}{\varrho} \left[\frac{12(1 - \beta)(\Gamma - 1) + \beta}{\Gamma - 1} d \ln T - 3(1 - \beta) d \ln \varrho \right] \\
&= \frac{p}{\varrho} \left[\frac{4 - 3\beta}{\Gamma_3 - 1} d \ln T - 3(1 - \beta) d \ln \varrho \right] \\
&= C_V T d \ln T - 3(1 - \beta) \frac{p}{\varrho} d \ln \varrho. \tag{10.110}
\end{aligned}$$

From the first law of thermodynamics we obtain

$$d\mathcal{E} = T ds - p d\frac{1}{\varrho} = T ds + \frac{p}{\varrho} d \ln \varrho \tag{10.111}$$

and therefore for the entropy variation

$$\begin{aligned}
T ds &= C_V T d \ln T - (4 - 3\beta) \frac{p}{\varrho} d \ln \varrho \\
&= C_V T [d \ln T - (\Gamma_3 - 1) d \ln \varrho], \tag{10.112}
\end{aligned}$$

where

$$\Gamma_3 - 1 \equiv \left(\frac{d \ln T}{d \ln \varrho} \right)_{\text{ad}}. \tag{10.113}$$

This gives the specific heat C_V

$$C_V = \frac{\mathcal{R}}{\mu(\Gamma - 1)} \left(\frac{12(1 - \beta)(\Gamma - 1) + \beta}{\beta} \right) = \frac{4 - 3\beta}{\Gamma_3 - 1} \frac{p}{\varrho T} \tag{10.114}$$

and the corresponding adiabatic index Γ_3

$$\Gamma_3 - 1 = \frac{(4 - 3\beta)(\Gamma - 1)}{12(1 - \beta)(\Gamma - 1) + \beta}. \tag{10.115}$$

In the case of gas pressure dominance we have $\Gamma_3 = \Gamma$. These equations are completely analogous to the corresponding equations in stellar structure.

With these relations one can derive the **evolution equation for the temperature** in the disk mid-plane, $T(t, R)$,

$$\begin{aligned}
\frac{\partial T}{\partial t} + V_R \frac{\partial T}{\partial R} - \frac{1}{C_V \Sigma} \frac{\partial}{R \partial R} \left[\frac{8acT^3 H}{3\kappa \Sigma} \frac{R \partial T}{\partial R} \right] \\
= \frac{(\Gamma_3 - 1)}{4 - 3\beta} \frac{T}{R} \frac{\partial(RV_R)}{\partial R} + \frac{Q^+}{C_V \Sigma} - \frac{Q^-}{C_V \Sigma}. \tag{10.116}
\end{aligned}$$

This is a diffusion-type equation for the temperature evolution with sources given by the compression of the flow and viscous heating, Q^+ , as well as by radiative losses Q^- . In most applications, the advective term and radial diffusion are neglected. This approximation is justified for standard disks, but it no longer holds for general accretion. In particular near the horizon extreme radial temperature gradients occur so that also heat conduction is an important contribution (this just changes the heat capacity in the diffusion term).

Cooling

The cooling function Q^- depends on the physical processes (Bremsstrahlung, cyclotron cooling, optically thick emission or Comptonization). For Bremsstrahlung one can write (in cgs units)

$$Q_{\text{Brems}}^- = 1.24 \times 10^{21} \text{ erg s}^{-1} \text{ cm}^{-2} H Q^2 \sqrt{T}. \quad (10.117)$$

For optically thick emission from the surface of the disk

$$Q^- = \frac{16acT^4}{3\tau} \quad (10.118)$$

with the optical depth $\tau = \Sigma k_{\text{eff}}/2$ of the disk. Similar expressions can be found for cyclotron cooling.

Turbulence

The question of the viscosity ν is of central importance for accretion disks. Since the time of Shakura and Sunyaev [363], most people use a phenomenological ansatz

$$\nu = \frac{2}{3} \alpha_{\text{SS}} c_S H, \quad (10.119)$$

with $\alpha_{\text{SS}} < 1$ as the turbulence parameter. Typical values found in magnetic turbulence are $\alpha_{\text{SS}} \simeq 0.01$.

Abramowicz et al. [22] first obtained a unified $\Sigma - \dot{M}$ picture for stationary accretion flows at a fixed radius R in the case of low viscosity, which includes four classes of solutions, namely SSDs, Shapiro–Lightman–Eardley (SLE) disks, slim disks and ADAFs. The solid lines in Fig. 10.10 represent thermal equilibrium solutions, i.e. stationary solutions of the 1D disk equations with $Q^+ = Q^-$. In Fig. 10.10, the right S-shaped curve consists of three branches, of which the lower one is for gas-pressure-supported SSDs, the middle one for radiation-pressure-supported SSDs, and the upper one for slim disks; while the left curve consists of two branches, of which the lower one is for SLE disks, and the upper one for ADAFs. The unstable branches are those which have $Q^+ > Q^-$ above and $Q^+ < Q^-$ below, while the stable branches are just opposite. Thus gas-pressure-supported SSDs, slim disks and ADAFs are thermally stable, whereas radiation-pressure-supported SSDs and SLE disks are thermally unstable.

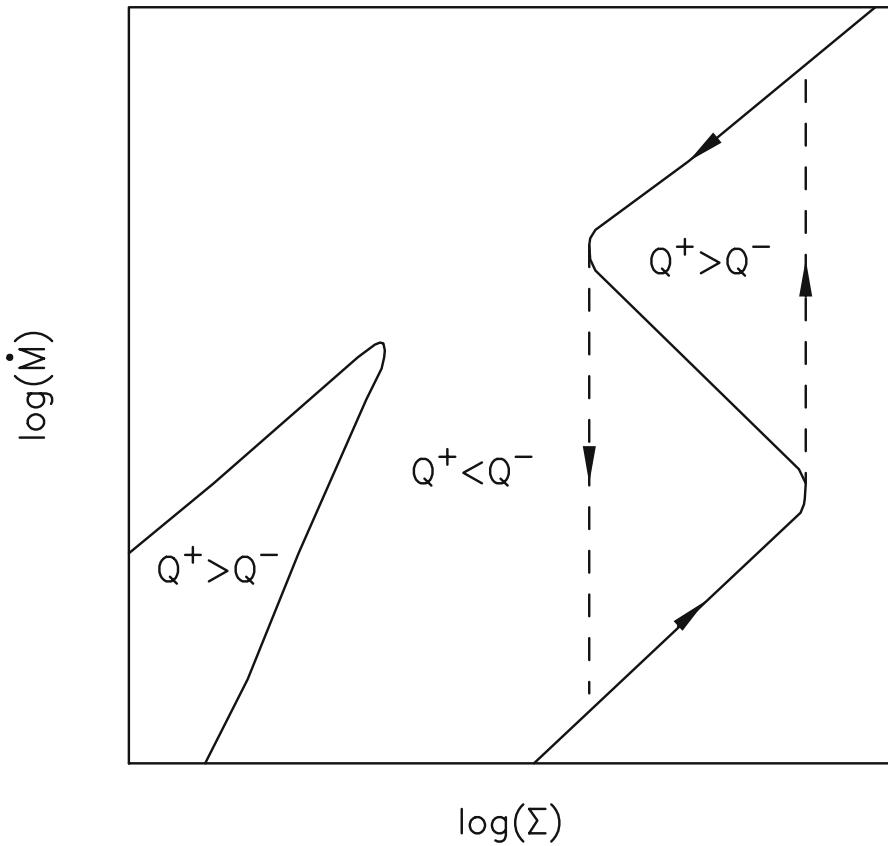


Fig. 10.10. Optically thin (*left branches*) and optically thick states (*right branches*) for accretion onto compact objects. Two types of $\log(\Sigma) - \log(\dot{M})$ branches exist for accretion flows at a fixed radius. Besides the standard models (SSD, *lower right branch*), two different branches appear for extremely low accretion rates (ADAF and the Shapiro–Lightman–Eardley (SLE) disk). The arrows show the behavior of the flow resulting from the thermal instability of a radiation-pressure-supported SSD

10.3.3 Standard Thin Disk Solutions (SSD)

When only radial advection is dominant, the above equations lead to the mass conservation

$$\frac{\partial \Sigma}{\partial t} + \frac{1}{R} \nabla_R (R \Sigma V_R) = 0, \quad (10.120)$$

and conservation of the specific angular momentum $j = R^2 \Omega$

$$\frac{\partial(\Sigma j)}{\partial t} + \frac{1}{R} \nabla_R (R \Sigma V_R j) = \frac{1}{R} \partial_R (R^2 W_{R\phi}). \quad (10.121)$$

Under the condition that $\partial_t \Omega = 0$, one can combine these two equations for Keplerian disks

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{R} \partial_R \left[\sqrt{R} \partial_R [\nu \sqrt{R} \Sigma] \right]. \quad (10.122)$$

In general, this is a nonlinear diffusion equation for $\Sigma(t, R)$, since ν can be a local function of the disk properties. This equation can be solved analytically for the case $\nu = \text{const}$

$$\Sigma(\tau, x) = \frac{m}{\pi R_0^2} \frac{1}{\tau x^{1/4}} \exp \left[-\frac{1+x^2}{\tau} \right] I_{1/4}(2x/\tau), \quad (10.123)$$

where $I_{1/4}(x)$ is the modified Bessel function. The initial condition is a ring of matter of mass m at radius R_0 , $x = R/R_0$ and $\tau = 12\nu t/R_0^2$ are dimensionless radius and time variables.

Stationary and geometrically thin disks are obtained under the following assumptions:

- **Mass conservation**

$$\dot{M} = 2\pi R V_R \Sigma = \text{const} \quad (10.124)$$

- **Almost Keplerian orbits**, i.e. $|V_R| \ll V_\phi$, $\Omega = \Omega_K$. This assumption has been confirmed in MRI simulations. The radial Euler equation is automatically satisfied.
- **Angular momentum conservation**: since there are no stable circular orbits inside the innermost stable orbit, this orbit is often assumed to be the inner boundary of the accretion disk with a torque-free condition, $W_{R\phi}(R_{\text{in}}) = 0$. The equation of angular momentum conservation can therefore be integrated

$$\frac{\dot{M}}{2\pi} (j - j_{\text{in}}) = -R^2 W_{R\phi}. \quad (10.125)$$

Since $j = \sqrt{GM_H R}$, this equation determines the torque on a ring in the disk, independent of the nature of the viscosity,

$$W_{R\phi} = \int_{-H}^H \pi_{R\phi} dz = \nu \Sigma R \frac{d\Omega}{dR} = -\frac{\dot{M}}{2\pi R^2} \sqrt{GM_H R} \left(1 - \sqrt{R_{\text{in}}/R} \right). \quad (10.126)$$

For Keplerian orbits, this relation can be written as

$$\nu \Sigma = \frac{\dot{M}}{2\pi} \left[1 - \sqrt{R_{\text{in}}/R} \right]. \quad (10.127)$$

- **Energy conservation**: The kinematic viscosity ν generates a dissipation of energy in the disk at a rate $D(R)$ per unit area and per unit time

$$Q^+(R) = D(R) = \frac{1}{2} \nu \Sigma (R \Omega')^2 = \frac{3GM_H \dot{M}}{8\pi R^3} \left[1 - \sqrt{R_{\text{in}}/R} \right]. \quad (10.128)$$

This expression is also independent of the nature of the viscosity and determines the total disk luminosity

$$L_{\text{disk}} = \int_{R_{\text{in}}}^{\infty} D(R) 2\pi R dR = \frac{1}{2} \frac{GM_H}{R_{\text{in}}}. \quad (10.129)$$

Half of the potential energy gained by moving inwards is dissipated away, the other half is stored in kinetic energy of the orbit. Matter just outside the boundary layer around the horizon still retains as kinetic energy half of the potential energy. For slowly rotating white dwarfs and neutron stars, this energy is emitted in the boundary layer, for black holes it is swallowed by the horizon. As we will discuss in Sect. 10.5, this inner boundary condition is violated by rapidly rotating black holes, where magnetic fields can interact with the frame-dragging effect.

- **Hydrostatic equilibrium:** Neglecting motions in the vertical direction, the vertical Euler equation reduces to the hydrostatic equilibrium

$$\frac{1}{\varrho} \frac{dP}{dz} = - \frac{GM_H}{R^2} \frac{Z}{R}. \quad (10.130)$$

This leads to the solution for the disk height H

$$H \simeq \frac{c_S}{\Omega_K}, \quad (10.131)$$

where c_S is the sound speed.

- **EoS:** The total pressure of the disk material is the sum of gas and radiation pressure

$$P = \frac{\varrho k_B T}{\mu m_p} + \frac{4\sigma_{SB}}{3c} T^4, \quad (10.132)$$

where k_B is the Boltzmann constant, μ the mean molecular weight, and σ_{SB} the Stefan–Boltzmann constant.

- **Radiative Transport:** The temperature T_c in the midplane of the disk is obtained by the energy equation relating the vertical energy flux to the rate of generation of thermal energy by viscous dissipation (or any other form of dissipation). When we equate the heat input per unit area, $Q^+(R)$, with the heat loss per unit area given by the radiative transport, we obtain for the central temperature

$$\frac{4\sigma_{SB}}{3\tau} T_c^4 = D(R) = \frac{3GM_H \dot{M}}{8\pi R^3} \left[1 - \sqrt{R_{\text{in}}/R} \right]. \quad (10.133)$$

τ is the optical depth of the disk defined through

$$\tau = \kappa_R \varrho H = \kappa_R \Sigma, \quad (10.134)$$

with $\kappa_R = \kappa_R(\varrho, T)$ as the total Rosseland mean opacity. For hot disks around compact objects, the main contributor to the opacity are free–free transitions and Thomson scattering. The optical thick condition for standard disks [363] implies that $\tau \gg 1$. If $\tau \leq 1$, radiation escapes directly, and the above estimate for the central temperature is no longer valid.

- **Turbulent viscosity:** The standard model of disks is based on the so-called α prescription of Shakura and Sunyaev [363]. Here, it is assumed that the dominant process for redistribution of angular momentum is turbulent viscosity, where the effective kinematic viscosity of a turbulent process is given by $\nu \simeq L_T v_T$ with L_T as the size and v_T as the velocity of the largest eddies in the flow. In accretion disks, we can assume that the scale-height of the eddies is less than the disk thickness H , and turbulence is subsonic. Consequently,

$$\nu = \alpha_{\text{SS}} c_S H, \quad (10.135)$$

with $\alpha_{\text{SS}} \leq 1$. This parametrization has been confirmed still to be useful, even in cases where turbulence is excited by magnetic instabilities (MRI), where one typically finds $\alpha_{\text{SS}} \simeq 0.1\text{--}0.01$ [74].

Solving the above listed set of equations, one obtains then the central density $\varrho_c(R)$, the surface density $\Sigma(R)$, the central pressure $P_c(R)$, the disk height $H(R)$, the radial drift $V_R(R)$, the central temperature $T_c(R)$, and the optical depth $\tau(R)$ as functions of the parameters M , \dot{M} and α . For fixed values of M and \dot{M} , the disk around neutron stars and black holes can be divided into three distinct regions:

- An **outer disk**, at large radii, in which gas pressure dominates over radiation pressure and in which the opacity is mainly due to free–free processes.
- A **middle region**, in which still gas pressure dominates over radiation pressure, but the opacity is mainly due to electron scattering.
- An **inner region**, near the surface of the compact object, in which radiation pressure dominates over gas pressure, and the opacity is still dominated by electron scattering. Since the optical depth in this inner part is not extremely high, Comptonization of soft disk photons by the hot electrons is a major source for SEDs.

Spectral Energy Distributions (SED) for SSD

The spectral energy distribution (SED) of the gravitational power released as electromagnetic radiation when matter accretes onto a compact object and its temporal variability are the main observed features for compact objects. The observed spectral energy distributions (SEDs) are far from universal. In standard disks, we would expect a multicolor black-body spectrum, determined by the temperature given by $\sigma_{SB} T_{\text{eff}}^4(R) = D(R)$,

$$T_{\text{eff}}(R) = \left\{ \frac{3GM\dot{M}}{8\pi R^3 \sigma_{SB}} \left[1 - \sqrt{R_{\text{in}}/R} \right] \right\}^{1/4}. \quad (10.136)$$

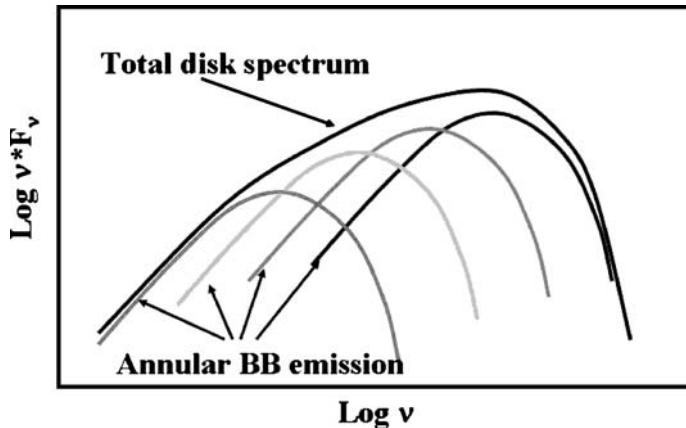


Fig. 10.11. Global spectrum of a standard disk consists of a superposition of black-bodies for each ring in the disk. The temperature of the rings steadily increases towards the central object

When the spectrum of each ring in the disk is approximated by a Planck spectrum, an observer at a distance d , whose line-of-sight makes an angle θ to the normal of the disk plane, would observe a spectral flux from the disk

$$F_v = \frac{4\pi h\nu^3 \cos \theta}{c^2 d^2} \int_{R_{\text{in}}}^{R_{\text{out}}} \frac{R dR}{\exp[h\nu/k_B T_*(R)] - 1}, \quad (10.137)$$

where R_{out} is the outer radius of the disk. In the integrated spectrum we can distinguish between the following branches (Fig. 10.11)

- $F_v \propto \nu^2$ for $h\nu \ll k_B T_{\text{out}}$;
- $F_v \propto \nu^{1/3}$ for $k_B T_{\text{out}} \ll h\nu \ll k_B T_{\text{max}}$;
- Wien decay for $h\nu \gg k_B T_{\text{max}}$.

NLTE effects in the disk atmosphere will change the above multicolor spectrum (Fig. 10.12), which is only valid for high accretion rates. For accretion rates below a certain threshold, roughly $\dot{M} < 0.03\dot{M}_{\text{ED}}$, the optical depth in the inner disk drops below unity – the inner disk becomes optically thin and is heated up. Different accretion modes are however possible, with different radiative properties. The main goal of accretion theory is to distinguish and understand all the possible different modes of accretion, and classify the observed sources in terms of such modes. The spectra of accreting black holes clearly show different components, that are produced by different physical processes and originate from different regions of the accretion flow. One of those components is a quasithermal hump, produced by a geometrically thin, optically thick accretion. It is shown that such a model not only underestimates systematically the value of the inner disk radius, but that, when the accretion rate through the disk is allowed to change, the inner edge of the disk, as inferred from the multicolor model, appears to move, even when it is in fact fixed at the innermost

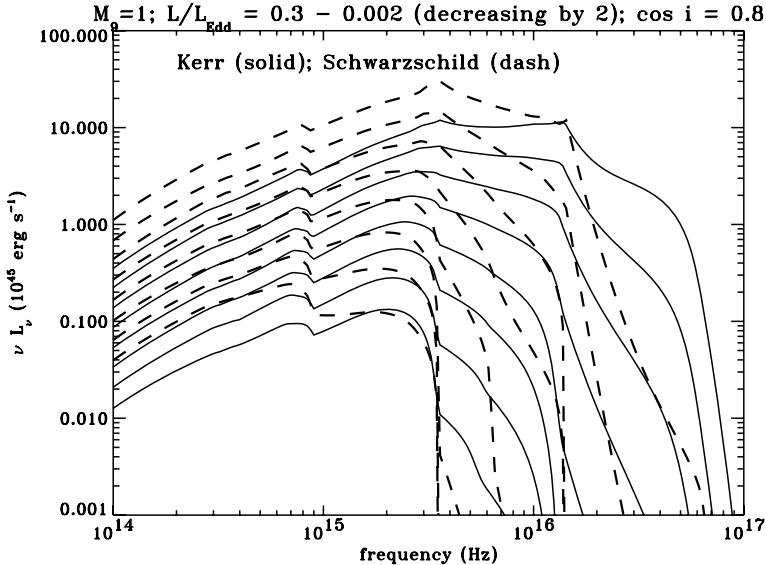


Fig. 10.12. Non-LTE disk spectra for quasars. A comparison of predicted spectral energy distributions for the sequence of models for a maximum rotating Kerr black hole (*solid line*), and for a Schwarzschild black hole (*dashed lines*), for models with $M = 10^9 M_\odot$ and various values of L/L_{Edd} . The values of the mass accretion rates are 2, 1, 1/2, 1/4, etc. for the Kerr hole, and the corresponding values for the Schwarzschild hole are 5.613 times larger. Figure adapted from Hubeny et al. [204]

stable orbit. Most accreting black holes, either of stellar mass or supermassive, when observed with hard X-rays show also signs of a hot Comptonizing component in the flow, the so-called corona.

Relativistic effects will also influence the form of the observed spectra. Li et al. [253] have applied a ray-tracing technique to compute the observed spectrum of a thin accretion disk around a Kerr black hole (relevant for the high state). They include all relativistic effects such as frame-dragging, Doppler boost, gravitational redshift, and bending of light by the gravity of the black hole (see Sect. 8.4.4). Self-irradiation of the disk as a result of light deflection is also included in these models. Assuming that the disk emission is locally black-body, the observed spectrum depends on the spin of the black hole, the inclination of the disk, and the torque at the inner edge of the disk. With this method, the spin parameters of the black holes in 4U1543–47 and GRO J1655–40 are estimated to be $a/M \simeq 0.6$ and $\simeq 0.6$ –0.7, respectively, or even somewhat higher, if color effects are included in the fitting procedure. These spectra will be modified by non-LTE effects due to Compton scattering. Models of this type applicable to the high/soft state of black hole X-ray binaries are discussed by Davis et al. [130]. For a given source, these models predict that the luminosity in the high/soft state should approximately scale with the fourth power of the empirically inferred maximum temperature, but with a slight hardening at high luminosities, in good agreement with observations.

10.3.4 Advection-Dominated Flows (ADAF)

The Advection-Dominated Accretion Flow model (ADAF) originates as an alternative solution to the same hydrodynamical equations of viscous differentially rotating flows that were solved in the standard disk model [211, 307, 308]. The ADAF model is a solution that works with low, sub-Eddington accretion rates. In this solution, the accreting gas has a very low density which implies that the flow is optically thin and therefore unable to cool efficiently within the accretion time. The viscous energy is therefore stored in the gas as thermal energy instead of being radiated away, and is advected onto the central compact object. The gas temperature is therefore extremely high. This causes the the gas to swell so that $H \simeq R$ – ADAF models are geometrically thick hot accretion disks. Since electrons can cool efficiently, but are not coupled tightly to the ions, ADAF models can evolve towards two-temperature flows, where the ions are much hotter than the electrons.

A low/hard state, characterized by a power-law spectrum with photon index of 1.4–1.8 and extending to 100 keV, is observed both in transient systems (X-ray

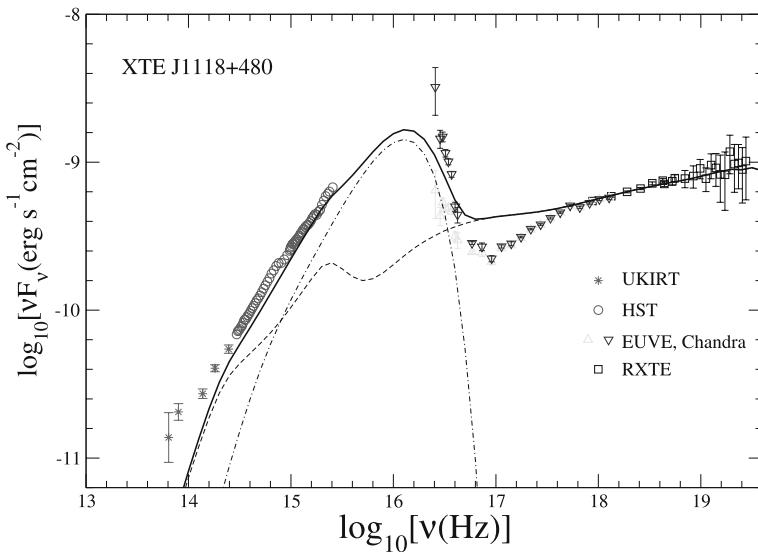


Fig. 10.13. Spectrum of a stellar black hole transient system. The absence of any soft black-body-like component in the X-ray band of the binary black hole source RXTE J1118+480 implies the existence of an extended hot optically thin region, with the optically thick cool disk truncated at some radius $R_{\text{tr}} \geq 50 R_S$. The *dashed-dot curve* represents the emission from a truncated accretion disk. Excess emission occurs in the infrared which extends down to the radio region. It has been associated with jet emission. The absence of the inner disk black-body component also underscores the requirement that the seed photons for thermal Comptonization (*dashed curve*) be produced locally in the hot flow, e.g. via synchrotron radiation. In the meantime, the mass of the black hole in this transient system has been determined to be $8.5 M_\odot$, the orbital period is 4.1 hours and the system is at a distance of 1.7 kpc [168]. Figure adapted from Yuan et al. [430]

novae, e.g. GRS 1124–68, GRO J0422+32) and persistent systems (e.g. Cyg X-1 and LMC X-3). It is generally assumed that these spectra are produced by thermal Comptonization of seed photons in the vicinity of an accreting black hole. Both Chandra and XMM are capable of observing soft X-rays down to 0.2 keV with good energy resolution. Esin et al. [148], and later Yuan et al. [430] present detailed modelling of the combined near-simultaneous HST, EUVE, Chandra and RXTE observations of the source RXTE J1118+480 (Fig. 10.13). A spectral model based on an advection-dominated accretion flow surrounded by a truncated thin disk and including jet outflows seems to provide reasonable fits to the combined optical, UV and X-ray data.

We will not discuss these models in details in the 1D approximation, since these models represent exactly the situations which can now be simulated in MHD turbulent processes.

10.3.5 Super-Eddington Accretion

For Eddington and super-Eddington accretion flows, radiation energy density is important. These models are preferentially discussed in spherical polar coordinates (r, ζ, ϕ) , where ζ is the polar angle measured from the equatorial plane of the disk. We also assume that the gas flow is axisymmetric. In this coordinate system, the basic equations for mass density, momentum, gas energy ϵ and radiation energy E are written in conservative form

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \mathbf{v}) = 0 \quad (10.138)$$

$$\begin{aligned} \frac{\partial(\varrho v)}{\partial t} + \nabla \cdot (\varrho v \mathbf{v}) &= \varrho \left[\frac{w^2}{r} + \frac{v_\phi^2}{r} - \frac{GM}{(r - R_S)^2} \right] - \frac{\partial P}{\partial r} \\ &\quad + \nabla \cdot \mathbf{S}_r + \frac{1}{r} S_{rr} + f_r \end{aligned} \quad (10.139)$$

$$\begin{aligned} \frac{\partial(\varrho w)}{\partial t} + \nabla \cdot (r \varrho w \mathbf{v}) &= -\tan \zeta \varrho v_\phi^2 - \frac{\partial P}{\partial \zeta} \\ &\quad + \nabla \cdot (r \mathbf{S}_\zeta) + \tan \zeta S_{\phi\phi} + f_\zeta \end{aligned} \quad (10.140)$$

$$\frac{\partial(\varrho r \cos \zeta v_\phi)}{\partial t} + \nabla \cdot (\varrho r \cos \zeta v_\phi \mathbf{v}) = \nabla \cdot (r \cos \zeta \mathbf{S}_{\phi\phi}) \quad (10.141)$$

$$\frac{\partial(\varrho \epsilon)}{\partial t} + \nabla \cdot (\varrho \epsilon \mathbf{v}) = -P \nabla \cdot \mathbf{v} + \Phi - \Lambda \quad (10.142)$$

$$\frac{\partial E}{\partial t} + \nabla \cdot \mathbf{F} + \nabla \cdot (\mathbf{v} F + \mathbf{v} P_R) = \Lambda - \varrho \frac{\kappa + \sigma}{c} \mathbf{v} \cdot \mathbf{F}. \quad (10.143)$$

ϱ is the density, $\mathbf{v} = (v, w, v_\phi)$ are the three velocity components, and P_R is the radiative stress tensor. Radiative quantities are calculated in the comoving frame. The force density exerted by radiation is given by

$$\mathbf{f} = \varrho \frac{\kappa + \sigma}{c} \mathbf{F}, \quad (10.144)$$

where κ and σ denote the absorption and scattering coefficients. The quantity Λ describes cooling and heating of the gas

$$\Lambda = \varrho c \kappa (S_* - E), \quad (10.145)$$

where S_* is the source function. In local thermodynamic equilibrium, $S_* = aT^4$, where T is the gas temperature. To close the system, we use the flux-limited diffusion approximation

$$\mathbf{F} = -\frac{\lambda c}{\varrho(\kappa + \sigma)} \nabla E \quad (10.146)$$

and

$$P_R = E T_{\text{Edd}}, \quad (10.147)$$

with T_{Edd} as the Eddington tensor.

Under these high accretion rates, the dominant radiation pressure force in the interior part of the disk accelerates gas vertically to disk plane, and jets with velocities of 0.2–0.4c are formed along the rotational axis [316]. The initially anisotropic jet flow expands outwards and becomes gradually isotropic at larger distances. Such outflows could explain the jets observed in the binary source SS 433 and the outflows observed in high-redshift quasars.

10.3.6 Unified Models of Disk Accretion

The most broadly accepted view on the structure of accretion disks is that the character of the flow is mainly determined by the Eddington ratio $\dot{m} = \dot{M}/\dot{M}_{\text{Edd}}$. In high Eddington ratio objects, accretion proceeds through a cold optically thick disk, while for lower Eddington ratio the cold disk evaporates close to the black hole. Below a certain radius r_{tr} , depending on the accretion rate, the accretion flow proceeds through some form of optically thin hot flow (ADAF). A plausible geometry is shown in Fig. 10.14.

For accretion rates above a few percent Eddington, accretion disks on scale much beyond the marginal stable orbit are well modelled by standard disk models (Fig. 10.14). These disks vertically collapse due to efficient radiative cooling which is the origin of the observed UV bumps in quasars. These disks will however be truncated at some radius r_{tr} , whose position is dictated by the overall accretion rate and the nature of the turbulence which is responsible for the angular momentum transport [182]. Here the disk is transformed into a hot geometrically thick and optically thin disk (sometimes called ADAF, or hot accretion torus). In this part, energy is mostly advected inwards and disappears inside the horizon of the black hole. In this part, the hard X-rays of Cyg X-1, Seyfert galaxies and quasars are generated by Compton scattering of soft photons.

One can analyze the formation of the transition front within time-dependent one-dimensional models based on a pseudo-Newtonian description of the gravitational

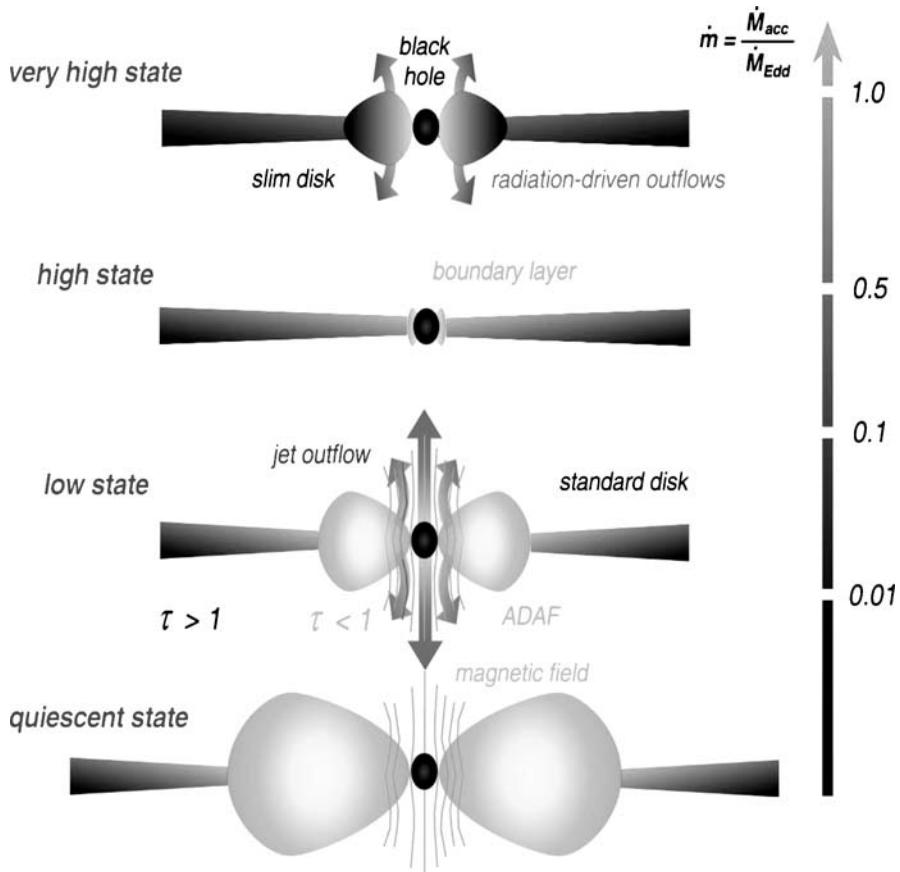


Fig. 10.14. Schematic sketch of the accretion flow in different spectral states as a function of the Eddington-scaled mass accretion rate \dot{m} in black hole systems (BBHs and galactic centers). Under high accretion rates, radiation pressure drives outflows, under low accretion rates, outflows are generated by magnetic coupling with the rotating black hole. Figure adapted from A. Müller [304] (LSW Heidelberg)

field of black holes [182], given in the previous section. This description is suitable to simulate the behavior of accretion onto nonrotating black holes. It turns out that the transition front is not stationary, but time-dependent. We found essentially two effects which drive the time-dependence of the transition radius: first, the transition radius has a tendency to drift to larger radii, the ADAF is growing and the SSD is shrinking. Secondly, also a rapid oscillatory motion of the transition radius is observed in numerical simulations [182].

Not much mass is probably lost from the standard disk. The growth of the black hole will therefore depend ultimately on the accretion processes occurring in the hot inner part. Many models show strong outflows under these conditions. One example

considered in some details is the case of nonradiative accretion flows driven by the magnetorotational instability, also called the Balbus–Hawley instability (De Villiers et al. [134, 135]; Brinkmann [88, 89]). The issue of the existence of super-Eddington accretion onto rotating black holes is therefore completely open. Super-Eddington accretion is, however, necessary to accelerate the growth of black holes in the early Universe in order to explain the bright quasars at redshift 6, where the Universe was less than one billion years old.

The time evolution of accretion states can be observed much easier for stellar black hole systems (see McClintock and Remillard [283]). Interestingly, the outer radius of the accretion disk is expected to be roughly one solar radius, and this corresponds to $\simeq 10^5$ gravitational radii, as in the case of extragalactic sources (e.g. M87). At all disk radii, the binding energy liberated by viscous dissipation is radiated locally and promptly and results in a gas temperature that increases radially inward reaching a maximum of $T \simeq 10^7$ K near the black hole. This picture is the standard thin disk model (SDD). This is at least correct for the high state spectra (HS). At lower mass accretion rates corresponding to several percent of the Eddington luminosity, a BHB usually enters the low/hard (LH) state and at very low accretion rates it reaches the quiescent state, which may be just an extreme example of the LH state (Fig. 10.14).

The hard X-ray spectrum is most plausibly explained as due to Comptonization of soft photons by the hot optically thin plasma in the inner torus.

The disks in NLS1 and quasars extend most probably down to the marginally stable orbit. This follows from the power-law character of their soft X-ray spectra and the change of the slope at the Lyman edge. In intermediate luminosity objects, like radio galaxies and normal Seyfert 1 galaxies, the disk most likely exists in the outer part of the flow, at distances of a few tens to a few hundreds of Schwarzschild radii. The argument here comes from the interpretation for double profiles of optical lines and the presence of relatively narrow iron lines (Reynolds and Nowak [341]).

10.3.7 Fundamental Time-Scales for Accreting Black Holes

If the accretion flow is roughly Keplerian, the most fundamental time-scale is the dynamical one given by the Keplerian frequency

$$t_{\text{dyn}} = 2\pi/\Omega_K(R) = \frac{2\pi R}{c} \sqrt{\frac{c^2 R}{GM_H}}. \quad (10.148)$$

We can conveniently express this time-scale in suitable units

$$t_{\text{dyn}} = 1.05 \text{ days } M_{H,9} (R/R_S)^{3/2}, \quad (10.149)$$

where $M_{H,9}$ denotes the mass of the black hole in units of 10^9 solar masses. Thermal time-scales are somewhat larger due to the turbulence parameter α

$$t_{\text{therm}} = t_{\text{dyn}}/\alpha. \quad (10.150)$$

When the transition radius changes with time, we need an estimate for the characteristic time-scale of the removal of the cold disk from a given radius. This removal occurs in a change of the accretion flow into an optically thin flow

$$t_{\text{evap}} = \frac{E}{\eta \dot{M} c^2} \quad (10.151)$$

with $E = \pi R^2 \Sigma kT/m_p$, which can be expressed as

$$t_{\text{evap}} \simeq 10 \text{ yrs } \dot{m}_{-1} M_{H,9} (R/10 R_S)^2. \quad (10.152)$$

For stellar mass black holes, this time-scale is of the order of seconds. In standard models, this time-scale can be identified with the viscous time-scale

$$t_{\text{evap}} = \frac{E}{F} = \frac{\Sigma kT}{\alpha P H \Omega_K} = t_{\text{visc}}, \quad (10.153)$$

where

$$t_{\text{visc}} = t_{\text{therm}} (R/H)^2. \quad (10.154)$$

In many radio-loud quasars, knots are ejected with a cyclic time-scale of a few years (in 3C 273 probably every year, in 3C 345 every fifth year). It is then tempting to relate this time-scale to the above evaporation time-scale. Since the mass in 3C 273 is probably similar to the black hole mass found in M87, where $M_H = 3 \times 10^9 M_\odot$, one year is close to the evaporation time for a radius of a few Schwarzschild radii. Since 3C 273 is a bright quasar, truncation must occur very near to the marginal stable orbit in order to explain the thermal UV bump of the spectrum. Contrary to this, the mass derived for 3C 120 from reverberation is $M_H = 3 \times 10^7 M_\odot$, and this source is probably in the low state (Marscher et al. [269]).

The X-ray emission observed from BH sources is highly erratic and can be analyzed in terms of power spectra. The accretion rate also determines the X-ray state of BH sources. In fact, high state and low state show different power density spectra (PDS), as exemplified by Cyg X-1 (Fig. 10.15).

Power spectra show typically a break from a slope -1 to -2 at some break frequency ν_{break} . In fact, all BH sources seem to have similar PDSs. In Fig. 10.16 we plot the break time-scales vs. Black hole masses for Cyg X-1 in the low and high state and for some Seyfert galaxies (for a break between a low frequency slope -1 and a high frequency slope of -2 (with data from McHardy et al. [284])). The NLS1s all lie above the high state line. If the break time-scale is associated with the position of the truncation radius and the same physics defines the characteristic time-scales in all black hole systems, then the movement of the truncation radius provides a natural explanation of Fig. 10.16. It is not clear at the moment, whether the position of the truncation radius is only a function of the accretion rate. The absence of broad iron emission lines in most Seyfert galaxies is a hint towards the correctness of this interpretation.

This scaling of the break frequency with the black hole mass demonstrates that the essential properties of black hole accretion are independent of the mass.

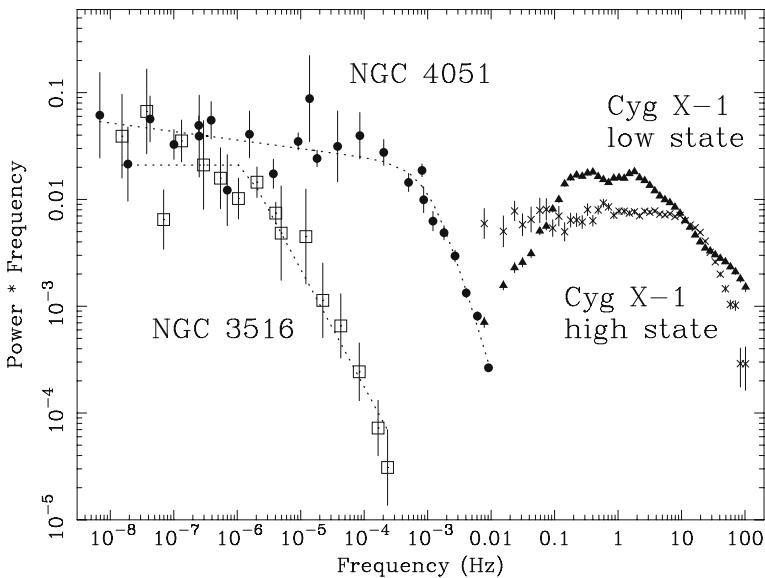


Fig. 10.15. PDSs of the Seyfert galaxy NGC 4051 (filled circles) and of Cyg X-1 in its low and high state. Frequency times power is shown here. Figure adapted from McHardy et al. [284]

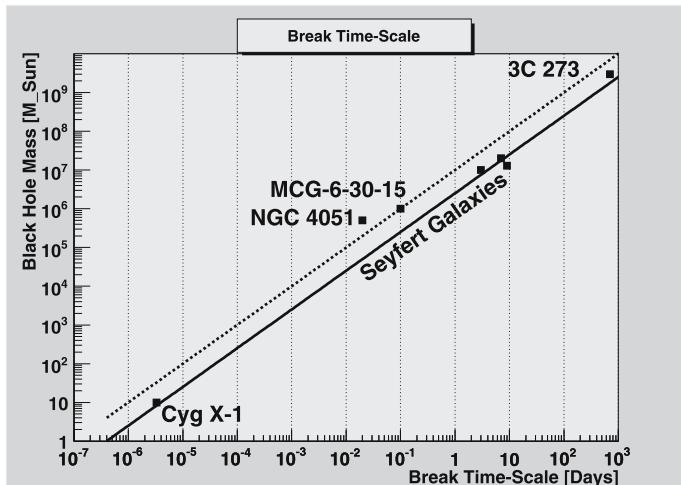


Fig. 10.16. Break time-scales for various AGN compared to Cyg X-1 in the low state. The Narrow Line Seyferts (NLS1) MCG-6-30-15 and NGC 4051 are shown together with broad-line Seyferts as *filled symbols*. The scaling of the break time with black hole mass is given for the low state (*solid line*) and the high state (*dotted line*). Data taken from McHardy et al. [284]

10.4 Relativistic MHD – Turbulent Accretion onto Black Holes

A powerful way to improve our understanding of astrophysical processes near compact objects is through accurate, large scale, three-dimensional numerical simulations. Nowadays, computational general relativistic astrophysics is an increasingly important field of research. In addition to the large amount of observational data gathered by high-energy X- and gamma-ray satellites such as Chandra, XMM–Newton, or INTEGRAL, and the new generation of gravitational wave detectors, the rapid increase in computing power through parallel supercomputers and the associated advance in software technologies now enables large scale numerical simulations in the framework of general relativity. However, the computational astrophysicist and the numerical relativist face a daunting task. In the most general case, the equations governing the dynamics of relativistic astrophysical systems are an intricate, coupled system of time-dependent partial differential equations, comprising the (general) relativistic (magneto-)hydrodynamic (MHD) equations and the Einstein gravitational field equations. In many cases, the number of equations must be augmented to account for nonadiabatic processes, e.g. radiative transfer or sophisticated micro-physics (realistic equations of state for nuclear matter, nuclear physics, magnetic fields, and so on).

10.4.1 From SRMHD to GRMHD

Magnetic fields play a crucial role in determining the evolution of many relativistic objects. Special relativistic MHD (SRMHD) is the foundation for any GRMHD

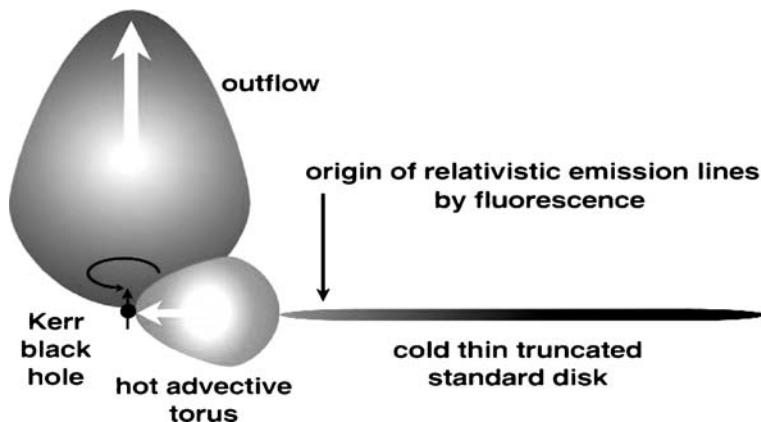


Fig. 10.17. Illustration of the main elements of modern accretion theory. Under the high accretion rates found in the early Universe, a standard accretion disk (SDD) is formed far away from the central black hole, which is truncated at a certain radius near the marginal stable orbit of the black hole. Here, a transition occurs towards an inner advection dominated disk which is of toroidal shape. The standard disk is the origin for the UV bump observed in quasars, while in the hot inner disk hard X-rays are produced by Compton scattering of soft photons. Figure provided by A. Müller [304] (LSW Heidelberg)

scheme, the transition to full GRMHD involves however some nontrivial problems. SRMHD schemes have been developed by van Putten [405], Komissarov [236], Bal-sara [53], Del Zanna et al. [132], Koldoba et al. [233], and Leismann et al. [251]. In recent years, numerical codes which evolve the general relativistic MHD equations (GRMHD) on a fixed Schwarzschild or Kerr black hole spacetime have been developed by Koide et al. [230, 232], Komissarov [238], De Villiers and Hawley [133], and Gammie et al. [166]. These codes have mainly been used to study the structure of accretion flows to rotating black holes, the Blandford–Znajek process in low-density magnetospheres [239, 240], and the formation of GRB jets [135]. A general approach to solve numerically the equations of GRMHD within the 3+1 split has been recently presented by Anton et al. [36], Baiotti et al. [46], and Duez et al. [138].

10.4.2 The Equations for GRMHD

For the following, we write the metric in the traditional 3+1 split

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ik}(dx^i + \beta^i dt)(dx^k + \beta^k dt). \quad (10.155)$$

As discussed in Sect. 2.8, a natural observer associated with the 3+1 splitting is the Eulerian observer with four-velocity \mathbf{n} perpendicular to the hypersurface of constant t . Its components are given by

$$n^\mu = (1, -\beta^i)/\alpha \quad (10.156)$$

and $n_\mu = (-\alpha, 0, 0, 0)$, respectively. The Eulerian observer is naturally extended by three spatial basis vectors that are tangent to the hypersurface $t = \text{const}$

$$\mathbf{e}_a = \{\mathbf{n}, \partial_i\}, \quad a = 0, \dots, 3. \quad (10.157)$$

In the one-component plasma limit, the equations of general relativistic MHD consist of particle conservation

$$\frac{1}{\sqrt{-g}} \partial_\mu [\sqrt{-g} \varrho_0 U^\mu] = 0, \quad (10.158)$$

expressed in terms of rest-mass density $\varrho_0 = mn$, and conservation of energy momentum

$$T^{\mu\nu}_{;\nu} = 0. \quad (10.159)$$

In the case of a one-component plasma, the stress–energy tensor consists of a perfect fluid part and the contributions from the Maxwell stresses (see Appendix E)

$$T^{\mu\nu} = \left(\varrho_0 + \epsilon + P + \frac{b^2}{4\pi} \right) U^\mu U^\nu + \left(P + \frac{b^2}{8\pi} \right) g^{\mu\nu} - \frac{1}{4\pi} b^\mu b^\nu, \quad (10.160)$$

where b^μ is the magnetic field as measured in plasma frame

$$b^\mu = -\frac{1}{2}\eta^{\mu\nu\rho\sigma} U_\nu F_{\rho\sigma}. \quad (10.161)$$

The three-velocity of the plasma as measured by the Eulerian observer is then given as

$$v^i = -\frac{h_\mu^i U^\mu}{n_\mu U^\mu}, \quad (10.162)$$

where $W = -n_\mu U^\mu = \alpha U^t$ is the Lorentz factor in the Eulerian frame and $h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu$ is the projector onto the hypersurface orthogonal to n^μ . The above expression can also be written as

$$v^i = \frac{U^i}{W} + \frac{\beta^i}{\alpha}. \quad (10.163)$$

A Conservative Formulation of GRMHD

Similar to the approach chosen to model pure hydrodynamical flows in Sect. 3.1, we shortly discuss the time evolution of magnetohydrodynamic fields based on a conservative schemes (details are given in Appendix E). Baryon number conservation gives

$$\partial_t[\sqrt{\gamma}D] + \partial_j[\sqrt{\gamma}Dv^j] = 0, \quad (10.164)$$

where $D = \varrho_0 \alpha U^t = \varrho_0 W$. Similar to the hydro case, we introduce the momentum fluxes measured by Eulerian observers

$$S_i = -n_\mu T_i^\mu = \alpha T_i^t = [\varrho_0 h + b^2/4\pi]W^2 v_i - \alpha b^t b_i/4\pi, \quad (10.165)$$

as well as the total energy density

$$\tau = n_\mu n_\nu T^{\mu\nu} = W^2 \varrho_0 h_* - P_T - \alpha^2 (b^t)^2 - D, \quad (10.166)$$

where $P_T = P + b^2/8\pi$ and $h_* = h + b^2/4\pi\varrho_0$.

The GRMHD equations have therefore the form of a hyperbolic system, similar to (3.32) (for a derivation, see Appendix E),

$$\frac{1}{\sqrt{-g}} \left(\frac{\partial[\sqrt{\gamma}\mathbf{U}]}{\partial t} + \frac{\partial[\sqrt{-g}\mathbf{F}^i]}{\partial x^i} \right) = \boldsymbol{\delta}, \quad (10.167)$$

which are obtained by combining the plasma equations with the induction equation (E.23). The state vector of GRMHD now consists of eight variables

$$\mathbf{U} = (D, S_i, \tau, B^i)^T, \quad (10.168)$$

explicitly given by the vector in the state space

$$\mathbf{U} = \begin{pmatrix} D \\ S_1 \\ S_2 \\ S_3 \\ \tau \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} = \begin{pmatrix} \varrho_0 W \\ (\varrho_0 h + b^2/4\pi)W^2 v_1 - \alpha b^t b_1/4\pi \\ (\varrho_0 h + b^2/4\pi)W^2 v_2 - \alpha b^t b_2/4\pi \\ (\varrho_0 h + b^2/4\pi)W^2 v_3 - \alpha b^t b_3/4\pi \\ (\varrho_0 h + b^2/4\pi)W^2 - P_T - \alpha^2(b^t)^2/4\pi - D \\ B^1 \\ B^2 \\ B^3 \end{pmatrix}. \quad (10.169)$$

P_T is the total pressure given by gas pressure $P = (\Gamma - 1)\varrho_0\epsilon$ and magnetic pressure, as defined above. \mathbf{B} is the magnetic field as measured by the Eulerian observer n^μ . The second part in the momentum fluxes $S_i, b^t b_i/4\pi \propto (\mathbf{v} \cdot \mathbf{B}) B_i/4\pi$, represents the Poynting fluxes. Starting with the expressions for b^t, b^i and b^2 in the definition of \mathbf{S} one finds the following expression for the total momentum fluxes

$$\boxed{\mathbf{S} = [W^2 \varrho_0 h + \mathbf{B}^2/4\pi] \mathbf{v} - \frac{(\mathbf{v} \cdot \mathbf{B})}{4\pi} \mathbf{B}}. \quad (10.170)$$

This is an essential feature of relativistic MHD, which is missing in the conservative formulation of MHD in the Newtonian approximation (see Sect. 10.3). In relativistic MHD, only the sum of kinetic momentum and Poynting flux is conserved. Relativistic MHD flows therefore allow for an exchange of momentum between kinetic momentum and momentum carried by electromagnetic fields. In the Newtonian version, this is only the case for the energy equation: the Bernoulli equation tells us that the total energy consisting of kinetic energy, internal energy and Poynting energy is conserved. This fact has already been observed in the case of stationary flows (see, e.g. Camenzind [98]) and is the secret behind the formation of relativistic jets. Relativistic jets are always generated, when the Poynting flux exceeds the total kinetic momentum at the base of the jet.

The corresponding fluxes \mathbf{F} are now given by

$$\mathbf{F}^i = \begin{pmatrix} DV^i \\ S_1 V^i - b_1 B^i/4\pi W + P_T \delta_1^i \\ S_2 V^i - b_2 B^i/4\pi W + P_T \delta_2^i \\ S_3 V^i - b_3 B^i/4\pi W + P_T \delta_3^i \\ \tau V^i + P_T v^i - \alpha b^t B^i/W \\ B^1 V^i - B^i V^1 \\ B^2 V^i - B^i V^2 \\ B^3 V^i - B^i V^3 \end{pmatrix}, \quad (10.171)$$

where $V^i \equiv v^i - \beta^i/\alpha$. The energy-momentum tensor in the sources \mathcal{S} now includes both parts, plasma and electromagnetic fields

$$\delta = \begin{pmatrix} 0 \\ T^{\mu\nu}\partial_\mu g_{\nu 1} - \Gamma_{\nu\mu}^\varrho g_{\varrho 1} \\ T^{\mu\nu}\partial_\mu g_{\nu 2} - \Gamma_{\nu\mu}^\varrho g_{\varrho 2} \\ T^{\mu\nu}\partial_\mu g_{\nu 3} - \Gamma_{\nu\mu}^\varrho g_{\varrho 3} \\ \alpha(T^{\mu t}\partial_\mu \alpha - T^{\mu\nu}\Gamma_{\nu\mu}^t) \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (10.172)$$

In this way, the above system generalizes the system of special relativistic MHD [236] (see also Leismann et al. [251]). GRMHD is not only important for the investigation of magnetized accretion disks around compact objects, but in the future also for core collapse calculations in supernovae, the merger of magnetized binary neutron stars, formation of supermassive black holes, and the generation of gamma-ray bursts (GRBs) [138]. It can be coupled with the time evolution of the gravitational fields, using, e.g. the BSSN formulation for the gravitational field [61, 367].

In GRMHD, the importance of magnetic effects is expressed in terms of two parameters, the plasma beta, $\beta = 8\pi P/b^2$, and a second parameter that defines the ratio of magnetic energy density to that of the rest mass

$$\delta = \frac{b^2}{4\pi\varrho_0 c^2}. \quad (10.173)$$

For $\delta \gg h$, the magnetic energy flux, $\varrho_0\sigma W^2\mathbf{v}$, will be much larger than the material energy flux, $\varrho_0 h W^2\mathbf{v}$. The speed of sound waves, c_s , follows from

$$hc_s^2 = \frac{\partial P}{\partial \varrho_0}|_\epsilon + \frac{P}{\varrho_0} \frac{\partial P}{\partial \epsilon}|_{\varrho_0}. \quad (10.174)$$

For an adiabatic EoS we thus find

$$c_s = \sqrt{\frac{\Gamma P}{\varrho_0 h}}. \quad (10.175)$$

The Alfvén speed, c_A , follows from Anile [33]

$$c_A^2 = \frac{b^2}{4\pi\varrho_0 h_T} = \frac{\delta}{h + \delta}. \quad (10.176)$$

This shows that for highly magnetized plasmas, $\delta \gg h$, and the Alfvén speed approaches the speed of light.

In numerical simulations, we also need primitive variables $\mathbf{P} = (\varrho_0, v^i, \epsilon, B^i)$. Here, $v^i = U^i/U^t$ is the three-velocity. The functions $\mathbf{U}(\mathbf{P})$ and $\mathbf{F}(\mathbf{P})$ are analytic, but the inverse operations are usually not. Also, there is no simple expression for $\mathbf{F}(\mathbf{U})$.

10.4.3 Nonradiative Accretion onto Rotating Black Holes

The study of relativistic accretion and black hole astrophysics is a very active field of research, both theoretically and observationally. Observations are providing increasingly detailed quantitative information about the accretion flows that power such high-energy systems as X-ray binaries and active galactic nuclei. These observations have been modelled in some detail by a variety of accretion scenarios, but such models rely on unavoidable assumptions such as regular flow geometry and a simple, parameterized stress. Global numerical MHD simulations offer a way to investigate the basic physical dynamics of accretion flows without these assumptions.

In recent years, there has been a revitalization of interest in magnetic fields near black holes. It is now widely believed that magnetohydrodynamic turbulence driven by the magnetorotational instability drives the flow of matter and angular momentum through the accretion disk. With that development, the properties of the magnetic field near to the central black hole have been revisited by several authors. Treating the region within the radius of marginal stability as force-free, some authors have argued that the large-scale magnetic field near the black hole is always weak due to the relatively weak fields at which the MRI saturates. If true, the extraction of the black hole spin energy via the Blandford–Znajek process may then be insignificant. On the other hand, one can make the opposite assumption (i.e. a dynamically insignificant magnetic field frozen into the plasma) and argue that the strongly sheared flow within the radius of marginal stability will amplify the field to very high values relative to the thermal pressure. Such fields may be effective at extracting the rotational energy of the black hole and/or creating a flow of energy and angular momentum from within the radius of marginal stability to outer parts of the accretion disk.

In the following we discuss a few results based on fully three-dimensional general relativistic magnetohydrodynamic simulation codes that evolve time-dependent inflows into Kerr black holes.

The time evolution of the density distribution in a weakly magnetized torus around a rotating black hole with Kerr parameter $a/M = 0.5$ has been simulated with the GRMHD code HARMS [166] (similar results were obtained by De Villiers and Hawley [134, 135], Hirose et al. [202], and Krolik et al. [246]; see also [200]). The pressure maximum is at 12 gravitational radii. Superposed on this equilibrium is a purely poloidal magnetic field with plasma beta of 0.01. The orbital period at the pressure maximum is 264 light crossing times. The simulation runs to 2000 light crossing times, corresponding to 7.6 orbital periods at pressure maximum. It will be interesting to investigate the coupling between magnetic fields and the gravitomagnetic field of the black hole occurring near the ergosphere of the black hole, i.e. typically in a region smaller than three gravitational radii. This is also the region where the velocities in pseudo-Newtonian calculations are to be the most inaccurate. In relativity, speeds are limited to the speed of light, while in pseudo-Newtonian simulations they can easily exceed this causal limit. For this purpose extremely high-resolution simulations are needed and the singular Boyer–Lindquist coordinates must be avoided.

Krolik et al. [246] have analyzed the bound matter inside the initial pressure maximum, where the time-averaged motion of gas is inward and an accretion disk forms. They use the flows of mass, angular momentum, and energy in order to understand dynamics in this region. The sharp reduction in accretion rate with increasing black hole spin observed in these simulations is explained by a strongly spin-dependent outward flux of angular momentum conveyed electromagnetically; when $a/M > 0.9$, this flux can be comparable to the inward angular momentum flux carried by the matter. In all cases, there is outward electromagnetic angular momentum flux throughout the flow; in other words, contrary to the assertions of traditional accretion disk theory, there is in general no stress edge, no surface within which the stress is zero. These conclusions are still based on nonconservative algorithms, for which angular momentum conservation has not shown to be true. The retardation of accretion in the inner disk by electromagnetic torques also alters the radial distribution of surface density, an effect that may have consequences for observable properties such as Compton reflection. The net accreted angular momentum is sufficiently depressed by electromagnetic effects that in the most

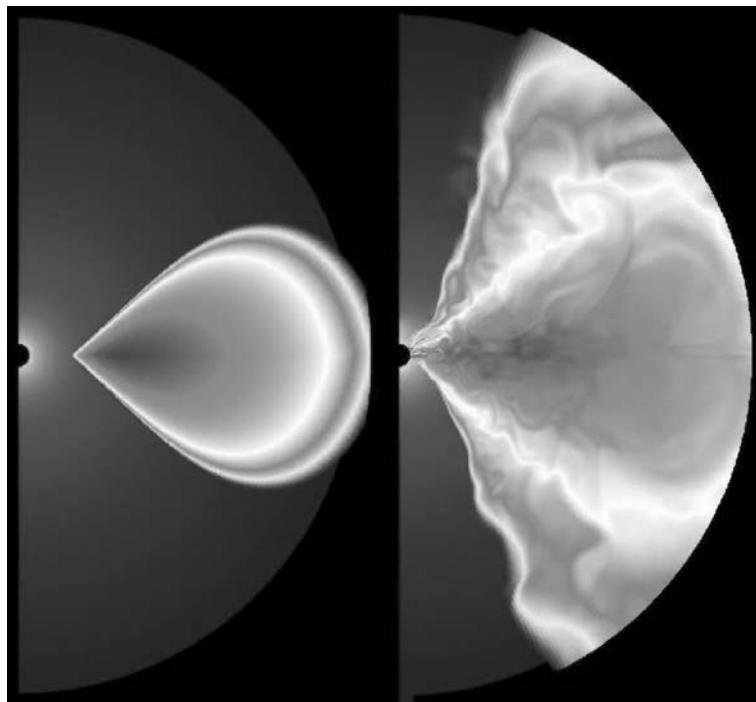


Fig. 10.18. Disk evolution near a rotating Kerr black hole with spin parameter $a/M_H = 0.5$ at $t = 0$ (*left*) and after 2000 light crossing times (*right*). The resolution is 300×300 grid points. After a few orbital periods at the pressure maximum the accretion becomes fully turbulent, as in the newtonian calculations. Figure adapted from Gammie et al. [166]

rapidly spinning black holes mass growth can lead to spin-down. Spinning black holes also lose energy by Poynting flux; this rate is also a strongly increasing function of black hole spin, rising to 10% or more of the rest-mass accretion rate at very high spin. As the black hole spins faster, the path of the Poynting flux changes from being predominantly within the accretion disk to predominantly within the funnel outflow.

10.5 Jets and the Ergosphere

The coupling of ordered magnetic fields within the ergosphere is a promising source for the generation of relativistic outflows from rapidly rotating compact objects (Fig. 10.19). The most compelling cases for a relativistic phenomenon are the ubiquitous jets in extragalactic radio sources associated with active galactic nu-

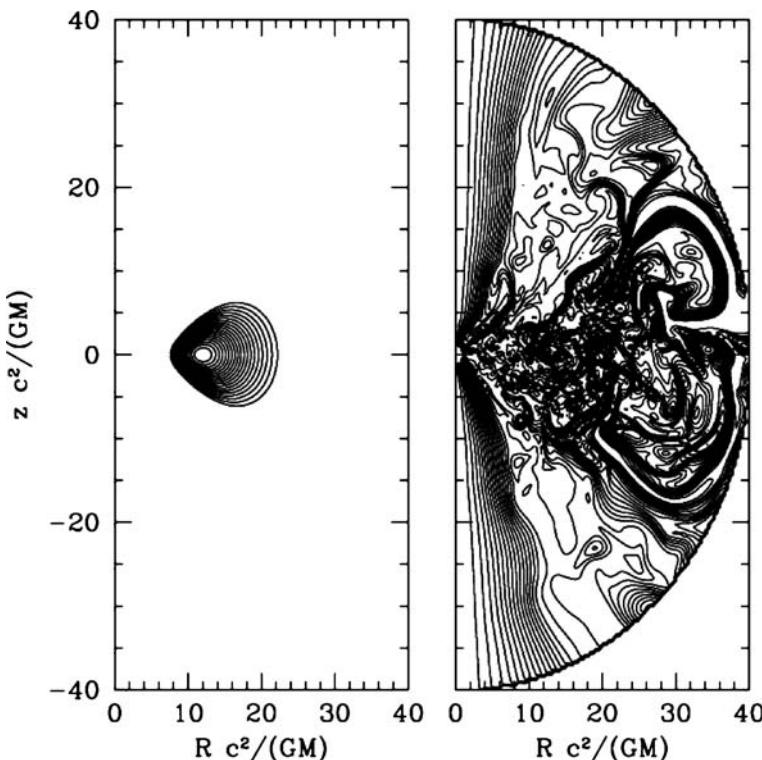


Fig. 10.19. Magnetic fields near rotating black holes. *Left:* initial field configuration is confined to the torus. *Right:* Field structure created by MRI. The field is highly turbulent in the disk and in coronal outflows, but strongly ordered in the polar funnel, which connects the ergosphere with the far zone. Cooling of the hot plasma will broaden the polar funnel. Figure adapted from Gammie et al. [166]

clei (see Fig. 10.22 and 10.25). In the commonly accepted standard model, flow velocities as large as 99% of the speed of light (and in some cases even beyond) are required to explain the apparent superluminal motion observed at parsec-scales in many of these sources. Models which have been proposed to explain the formation of relativistic jets involve accretion onto a compact central object, such as a neutron star or stellar mass black hole in the galactic microquasars GRS 1915+105 and GRO J1655–40, or a rotating supermassive black hole in an active galactic nucleus, which is fed by interstellar gas and gas from tidally disrupted stars.

At parsec scales, the jets, observed via their synchrotron and inverse Compton emission at radio frequencies with VLBI imaging, appear to be highly collimated with a bright spot (the core) at one end of the jet and a series of components which separate from the core, sometimes at superluminal speeds. In the standard model, these speeds are interpreted as a consequence of relativistic bulk motions in jets propagating at small angles to the line-of-sight with Lorentz factors up to 20 or more. Moving components in these jets, usually preceded by outbursts in emission at radio wavelengths, are interpreted in terms of travelling shock waves.

10.5.1 Jets as Outflows from the Ergospheric Region

One of the most spectacular phenomena associated with accretion onto black holes is the creation of powerful, highly relativistic jets. However, despite intense observational and theoretical study, the basic energy source of these relativistic jets remains unknown. Broadly speaking, there are two possibilities. Firstly, jets could be powered by the liberation of gravitational potential energy of accreting matter. If this is the case, the most likely scenario is the formation and subsequent focusing and acceleration of a magnetohydrodynamic (MHD) disk wind. While this mechanism has the appealing feature of potentially being universal to all accreting systems (and therefore allowing a unified model for jets from protostellar systems, accreting white dwarfs and accreting neutron stars as well as accreting black holes), it is not clear that such a disk wind can be accelerated to the highly relativistic velocities seen from many black hole systems. The alternative is that jets could be powered by the magnetic extraction of the spin energy of the central black hole using the mechanism described in the seminal paper by Blandford and Znajek [80]. The power extracted from a Kerr black hole with dimensionless spin parameter a_* threaded by a magnetic field of strength B_H (in the membrane paradigm sense) is

$$L_{\text{BZ}} \simeq \frac{1}{32} \omega_F^2 B_H^2 r_H^2 a_* c, \quad (10.177)$$

where r_H is the radius of the event horizon and $\omega_F^2 = \Omega_F(\Omega_H - \Omega_F)/\Omega_H^2$, with Ω_H and Ω_F being the angular velocities of the black hole and magnetic field lines, respectively. It is often argued (e.g. see [80]) that the magnetic field structure adjusts itself such that $\Omega_F = \Omega_H/2$, hence maximizing ω_F^2 to a value of 1/4. While the initial work of BZ was based on force-free black hole magnetospheres, the basic mechanism is seen to operate in the recent generation of fully relativistic MHD accretion disk simulation.

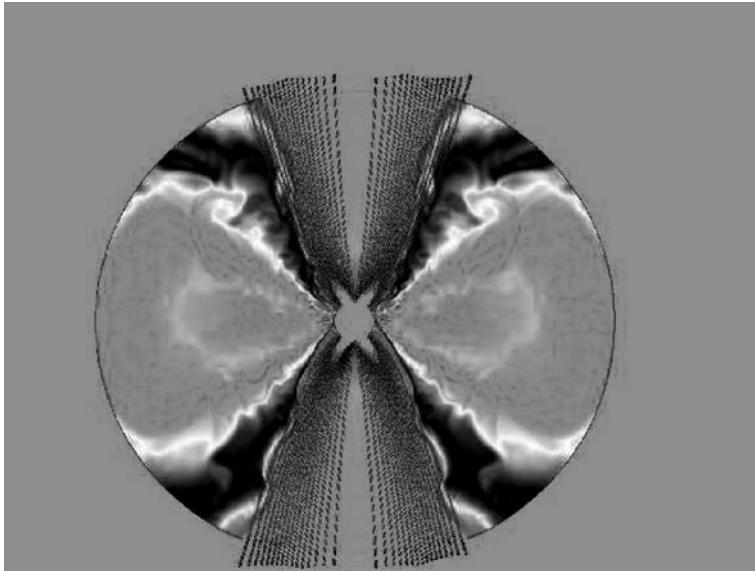


Fig. 10.20. Jets from the ergosphere. Turbulent accretion onto rotating black holes generates three different regions around the black hole. The hot turbulent disk is located near the equatorial plane with weak magnetic fields. Adjacent is a region of hot coronal outflows (disk winds), and near the polar region we find the polar MHD wind funnel which is completely magnetically dominated. Magnetic fields in the polar funnel are found to be well ordered. Figure adapted from McKinney [281]

In many studies, the plunge region of the black hole accretion disk has been neglected (see Fig. 10.18). This is the region of the disk within the radius of marginal stability, in which the accretion flow is undergoing rapid inwards acceleration (ultimately crossing the event horizon at the velocity of light as seen by a locally nonrotating observer). Unless the magnetic field is extremely strong, this is a region where inertial forces will dominate and the commonly employed force-free approximation will break down. This means that the actual field threading the plunge region would be very weak. However, it does not imply that the field threading the BH horizon, which is what counts for the BZ effect, is also weak. The field swept in by the plunge region would be cleaned into some well-ordered configuration threading the black hole (Fig. 10.20) and can be confined by the inertial forces of the plunging accretion flow even if it achieves strengths appreciably higher than the characteristic field strengths in the inner disk. Since the strength of the BZ mechanism depends on the square of the magnetic field, this enhancement could have major implications for the relative dominance of spin-energy extraction.

The jet structure that emerges from GRMHD simulations (Fig. 10.20) is primarily due to the internal evolution of the jet rather than by the interaction with the surrounding medium itself. In particular, these GRMHD-based results are insensitive to the two different models of the initial surrounding medium. One model

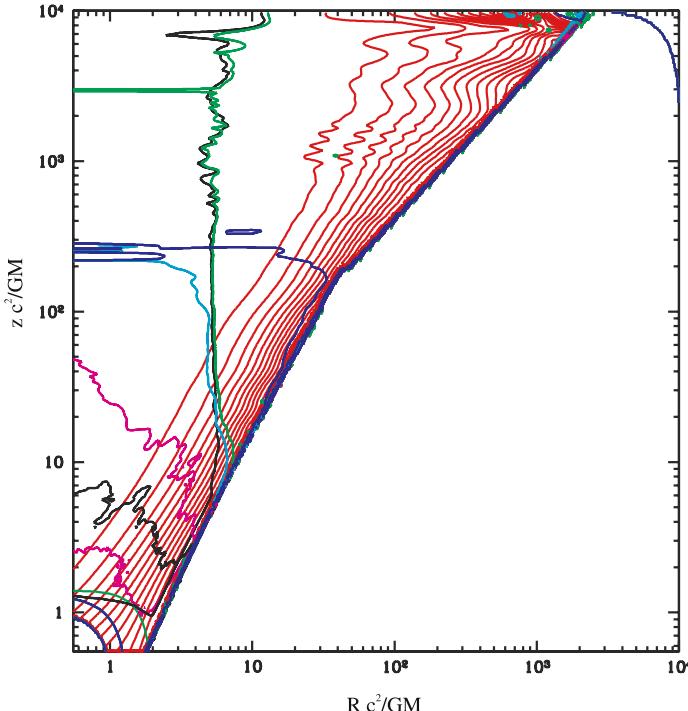
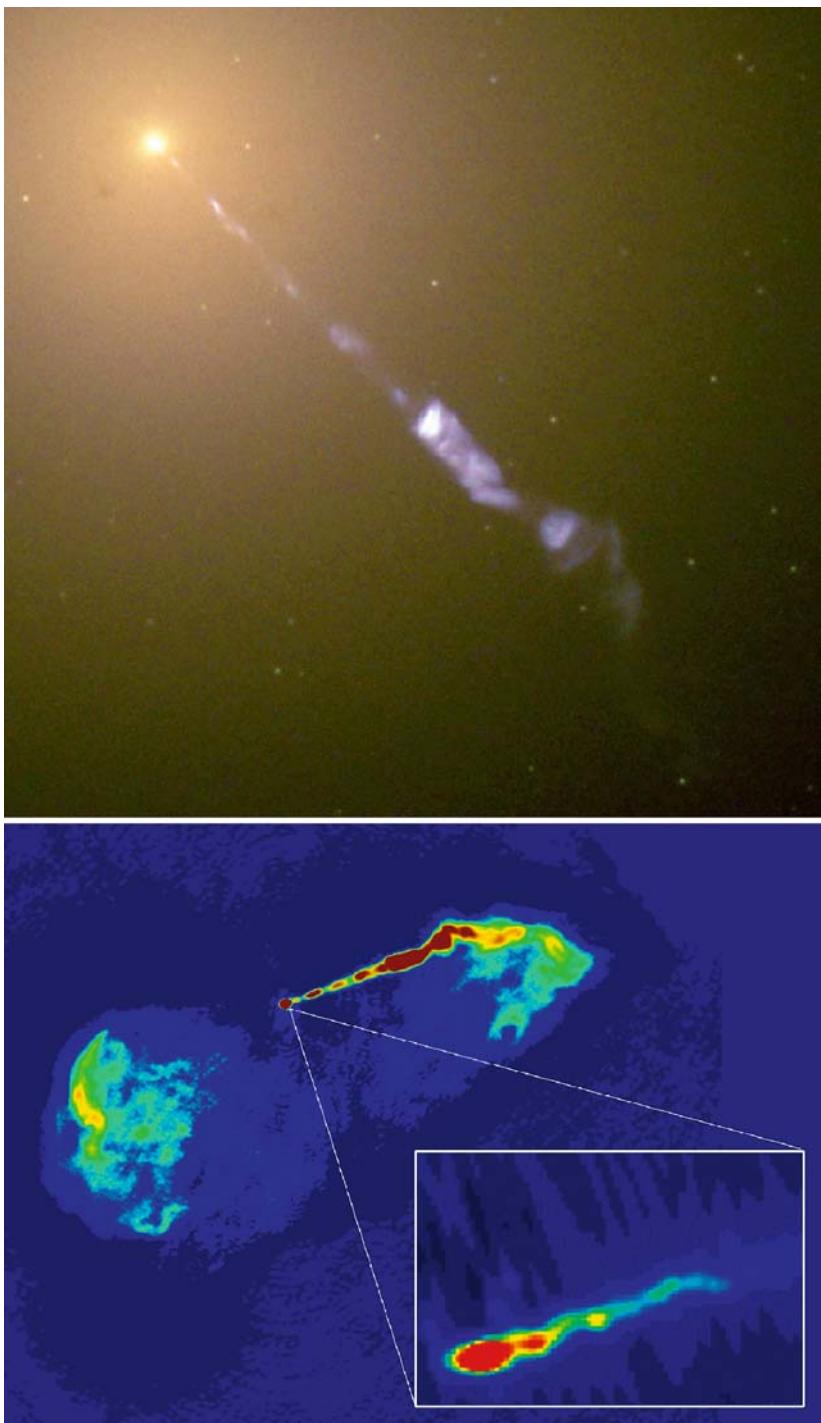


Fig. 10.21. Relativistic jets are driven by magnetic fields covering the ergosphere of rapidly rotating black holes. In a log–log plot, 45° lines correspond to lines of constant θ . Lines of constant spherical polar r are horizontal near the z -axis and vertical near the R -axis. The field lines are shown as red lines. From $r = 0$ outwards: Blue#1: horizon + ingoing-fast magnetosonic; Cyan#1: ingoing-Alfvén; Black#1: $|B^\phi| = |B^r|$; Green#1: ergosphere; Purple#1: ingoing-slow ; Black#2: stagnation surface where poloidal velocity $u^p = 0$; Purple#2: outgoing-slow ; Cyan#2: outgoing-Alfvén; Black#3: $|B^\phi| = |B^r|$ again ; Green#2: light cylinder; Blue#2: outgoing-fast. The disk and coronal regions have been truncated with a power-law cutoff for $r \leq 100 M$ and a conical cutoff for larger radii. Figure adapted from McKinney [281]

►

Fig. 10.22. The center of the Virgo galaxy Messier 87. A black hole with a mass of three billion solar masses is the origin of the jets visible from radio to X-rays in Messier 87. *Top:* In this HST image, the blue of the jet (synchrotron emission) contrasts with the yellow glow from the combined light of billions of unseen stars and the yellow, point-like globular clusters that make up this galaxy. The jet is produced by a 3-billion-solar-mass central object. *Bottom:* Radio image of M87, taken with the VLA, shows giant bubble-like structures where radio emission is thought to be powered by the jets subatomic particles coming from the central black hole. The false color corresponds to the intensity of the radio energy being emitted by the jet. *Inset:* The VSOP radio image of the region close to the black hole, where an extragalactic jet is formed into a narrow beam by magnetic fields. Image Credit: NASA and The Hubble Heritage Team (STScI/AURA) and VLA



considered by McKinney [281] is a surrounding infall of material and the other is an evacuated exterior region. The Poynting-dominated jet structure is negligibly broader or narrower in the evacuated case due to magnetic confinement. Beyond the outer Alfvén surface at $r \simeq 10 GM_H/c^2$, the jet becomes marginally unstable to current-driven instabilities. Such instabilities drive shocks in the jet that limit the efficiency of magnetic acceleration and collimation, but are vital for understanding the acceleration of nonthermal particles.

Figure 10.21 shows the characteristic structure of ergospheric jets for one polar axis as a result from long integration [281]. The inner Alfvén surface lies inside the ergosphere, as required to extract energy. The field lines follow nearly a power law until $r \sim 10^2 GM_H/c^2$. The stagnation surface, which divides inflow from outflow regions, is time-dependent, but stable. Clearly the transition to a supercritical (superfast) flow has occurred. After the fast surface, the field lines stretch out and oscillate around a conical asymptote. The fast surface near the polar axis is at $r \sim 250 M$, where for the other hemisphere it is at $r \sim 500 M$. This flow structure is exactly what one expects from stationary MHD. The outer Alfvén surface is located very near to the light cylinder, and this is responsible for relativistic outflow speeds.

Jets are a common outcome of accretion (Fig. 10.22), yet the observed jet properties, such as collimation and speed, are not uniform between systems. This is despite the fact that the basic physics (general relativistic magnetohydrodynamics (GRMHD)) to describe such systems is black hole mass-invariant. Thus, it is worthwhile determining the unifying, or minimum number of, pieces of physics that would explain most of the features of gamma-ray bursts (GRBs), X-ray binaries, and active galactic nuclei (AGN). To understand jet formation requires at least explaining the origin of the energy, composition, collimation, and Lorentz factor.

Primarily, one can distinguish two types of jets: (i) Poynting-dominated jets, which are typically dominated in energy flux by Poynting flux and dominated in mass by electron–positron pairs or electron–proton plasma for AGN and X-ray binaries, while probably dominated in mass by electron–proton plasmas for GRBs; and (ii) Poynting baryon jets with about equal Poynting flux and rest-mass flux and dominated in mass by baryons. The latter are sometimes referred to as coronal outflows due to their origin [285–287]. This two-component jet model is one key to understanding the diversity of jet observations. The Poynting-dominated jet is likely powered by the Blandford–Znajek effect, while the Poynting baryon jet is likely powered by both Blandford–Znajek power and the release of disk gravitational binding energy. Collimation of the polar Poynting-dominated jet may be due to transfield balance against the broader Poynting baryon jet or by self-collimating hoop stresses.

Jet Systems

Some black hole X-ray binaries have jets [155, 296], such as GRS 1915+105 with apparently superluminal motion ($\Gamma \sim 3$), but may have $\Gamma \sim 1.5$. Synchrotron radiation from the jet suggests the presence of a magnetized accretion disk. Observations of a broad, shifted, and asymmetric iron line from GRS 1915+105 is possible evi-

dence for a relativistic accretion disk, although this feature could also be produced by a jet component.

The standard paradigm is that relativistic jets from X-ray binaries are probably produced by the Blandford–Znajek effect. However, at least some black holes, such as GRS 1915+105, might have slowly rotating black holes. If this were correct, then another mechanism would be required to produce jets. Indeed, jets or outflows are produced from systems containing neutron stars, young stellar objects, supersoft X-ray white dwarfs, symbiotic white dwarfs, and even UV line-driven outflows from massive O stars.

AGN are observed to have jets with $\Gamma \leq 10$, or even $\Gamma \sim 30$. Some radio-quiet AGN show evidence of weak jets, which could be explained as a coronal outflow and not require a rapidly rotating black hole. Observations imply the existence of a two-component jet structure with a Poynting jet core and a dissipative surrounding component. The energy structure of the jet and wind are important in understanding the feedback effect that controls size of the black hole and may determine the $M-\sigma$ relation.

Neutron stars and black holes are associated with the most violent of post-Big-Bang events: supernovae and some gamma-ray bursts (GRBs) and probably some X-ray flashes (XRFs) [329]. Observations of a supernova light-curve in the afterglow suggest that at least some long-duration GRBs are probably associated with core-

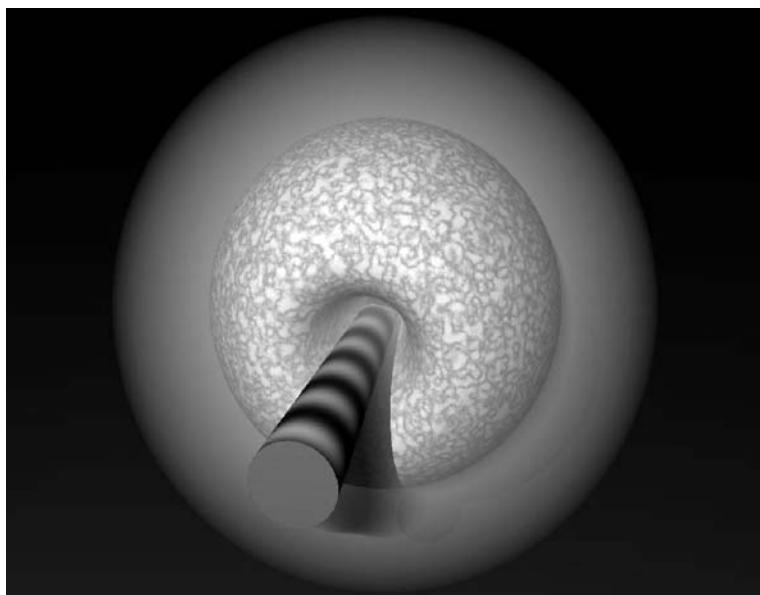


Fig. 10.23. AGN jets emerge from the immediate vicinity of black holes in the center of galaxies (red: stellar nucleus). The AGN jet is optically thin and emits nonthermal synchrotron and inverse Compton emission, when embedded into a bath of stellar and infrared photons from the dusty torus on the parsec-scale (cloudy medium)

collapse events (SN Ibc). Neutrino processes and magnetic fields are both important to understand core-collapse. In unravelling the mechanism by which core-collapse supernovae explode, the implementation of accurate neutrino transport has been realized to be critical to whether a supernova is produced in simulations. This has thus far been interpreted to imply that highly accurate neutrino transport physics is required, but this could also mean additional physics, such as a magnetic field, could play a significant role. Indeed, all core-collapse events may be powered by MHD processes rather than neutrino processes. Core-collapse involves shearing subject to the Balbus–Hawley instability as in accretion disks. All core-collapse explosions are significantly polarized, asymmetric, and often bi-polar indicating a strong role of rotation and a magnetic field.

Black hole accretion is the key source of energy for many GRB models. Collapsar type models suggest that a black hole forms during the core-collapse of some relatively rapidly rotating massive stars. The typical radius of the accretion disk likely determines the duration of long-duration GRBs. An accretion disk is also formed as a result of a neutron star or black hole collisions with another stellar object.

GRBs are believed to be the result of an ultrarelativistic jet. Indirect observational evidence of relativistic motion is suggested by afterglow achromatic light breaks and the “compactness problem” suggests GRB material must be ultrarelativistic with Lorentz factor $\Gamma \geq 100$ to emit the observed nonthermal gamma-rays (see, e.g. [329]). Direct observational evidence for relativistic motion comes from radio scintillation of the ISM and measurements of the afterglow emitting region from GRB030329.

Typical GRB jet models invoke either a hot neutrino-driven jet or a cold Poynting flux-dominated jet, while both allow for comparable amounts of the accretion energy to power the jet. A neutrino-driven jet derives its energy from neutrino annihilation from gravitational energy and the jet is thermally accelerated. However, strong outflows can be magnetically driven. In particular, black hole rotational energy can be extracted as a Poynting outflow [80].

10.5.2 From the Ergosphere to the Cluster Gas

Jets launched from the immediate vicinity of the black hole are collimated on the subparsec-scale in extragalactic sources and propagate then into the kiloparsec-scale, where they begin to interact with the surrounding cluster gas.

Jets and the Cluster Gas

The lobes of powerful radio galaxies are fuelled by two oppositely directed relativistic beams emitted from the nucleus. By comparing the power supplied by the beams with the power radiated by the radio components, one can show that only a small fraction of the power supplied can be radiated away from the tip of the beam, and that most of the energy must be deposited in a cavity surrounding the beam (Fig. 10.24).

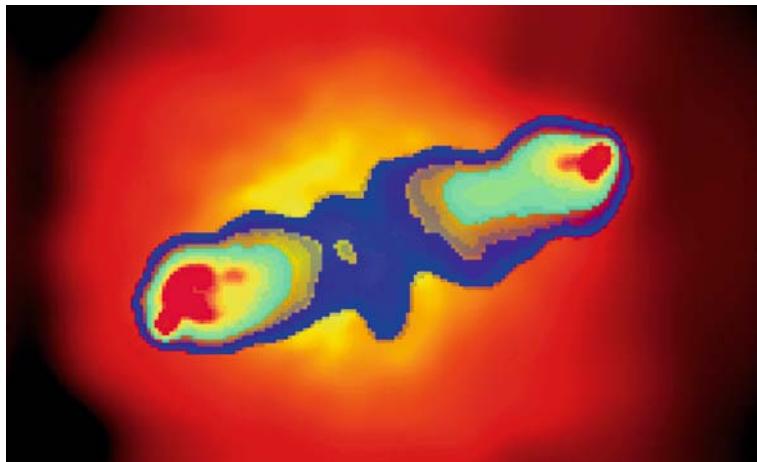


Fig. 10.24. Cygnus A and its cluster gas. The *blueish* colors represent low-frequency (330 MHz) synchrotron emission from the shocked beam plasma which is distributed in the cocoon around the jets on the kiloparsec-scale. The *reddish* and *yellow* colors represent X-ray emission from the cluster gas, which is strongly distorted in the central part (*yellow filaments*). The jets drive a weak elliptical bow-shock into the cluster gas. Image: superposition of VLA measurements at 330 MHz and according to a Chandra image

Using NASA's Chandra X-ray Observatory, astronomers have found a giant football-shaped cavity within X-ray emitting hot gas surrounding the galaxy Cygnus A. The cavity in the hot gas has been created by two powerful jets emitted from the central black hole region in the nucleus of Cygnus A. Hot gas is steadily being piled up around the cavity as it continuously expands, creating a bright rim of X-ray emission. The jets themselves terminate in radio and X-ray emitting **hot spots** some 70 kiloparsecs from the center of the galaxy.

Large-Scale Jets and High-Energy Emission

Large-scale jets extending to hundreds of kiloparsec distances from the quasar nucleus often radiate their power predominantly in X-rays. This has been revealed by surveys with the Chandra X-ray Observatory [270], following its unexpected discovery of an X-ray jet in the quasar PKS 0637–752 at redshift $z = 0.65$. Despite extensive observational and theoretical work, the dominant X-ray emission mechanism operating in quasar jets remains unsettled. Neither naive single-component synchrotron nor synchrotron-self-Compton models

can fully explain the spectral energy distributions (SEDs) at radio, optical and X-ray wavelengths. A widely discussed hypothesis for the strong X-ray emission is *relativistically enhanced inverse Compton* (IC) scattering off the cosmic microwave background (CMB) photons. In this model, the bulk velocity of the jet is assumed to be highly relativistic all the way to nearly megaparsec distances, with a Doppler factor of $D \sim 10$.

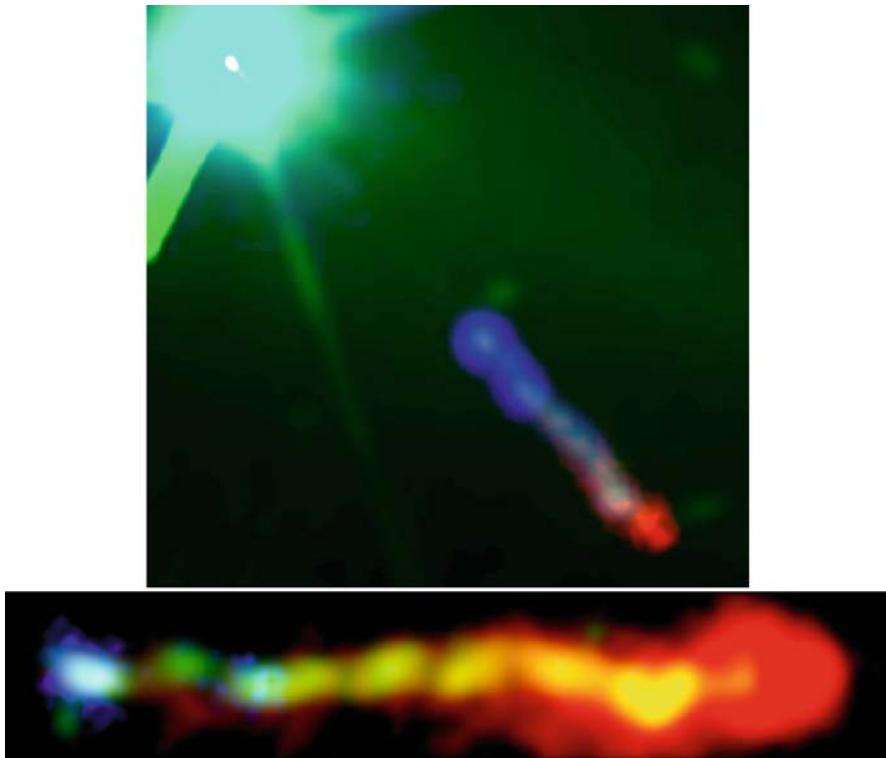


Fig. 10.25. The quasar 3C273 at redshift 0.158 and its jet. *Top:* False-color composite showing the relation between the quasar 3C273 (*top left*; the quasar is really just a very small and bright source, the fuzz apparently surrounding it is an artifact that appears when taking a picture of a very bright source) and the jet. The color coding is the same as in the image below. *Bottom:* False-color composite of 3C273's jet, showing in which wavelength region the emission peaks: X-rays (observed with Chandra) in *blue*, optical light (observed with HST) in *green*, radio waves (observed with the VLA) in *red*. *Yellow* indicates that both optical and radio emission are strong. The jet behaves differently at these three wavelengths: the X-ray image becomes fainter farther from the nucleus, the radio image gets brighter, and the optical remains largely constant, but fragmented into many knots. Details can be found in Uchiyama et al. [404]. Image Credit: HST, Spitzer and Chandra Observatory NASA/CXC/SAO

Alternative scenarios for producing X-ray emission at such large distances from the central engine include a variety of nonconventional synchrotron models. The X-ray emission could be explained by a spectral hump in a synchrotron spectrum as a result of reduced IC cooling in the Klein–Nishina regime provided that a Doppler factor of the jet is large. Alternatively, a second synchrotron component responsible for the X-rays could be formed by turbulent acceleration in the shear boundary layer or by hypothetical high-energy neutral beams (neutrons and gamma-rays) from the central engine. Finally, synchrotron radiation could also be produced by very high-energy protons, $E_p \geq 10^{18}$ eV.

Uchiyama et al. [404] have presented results from Spitzer IRAC observation of the 3C 273 jet at wavelengths 3.6 and $5.8\ \mu\text{m}$, combined with the photometry with the VLA radio, HST optical, and Chandra X-ray data (Fig. 10.25). This multiwavelength analysis led to the conclusions that the flat optical emission in the X-ray-dominated knots originates in the high-energy power-law component, which also accounts for the X-ray emission. The agreement between the optical-X-ray slope and the X-ray spectral index throughout the jet supports this picture. On the other hand, the radio to infrared spectra can be expressed by a power law with an exponential cutoff at $\sim 5 \times 10^{13}\ \text{Hz}$. The two distinct radiation components, namely low-energy (radio-infrared) and high-energy (optical-X-ray) emission, have similar power of $\sim 10^{44}\ \text{erg s}^{-1}$ for the entire jet volume (without a beaming correction); the power in the second component becomes noticeably higher, if its peak position is located far beyond the X-ray domain. The relative importance of the two components changes along the jet (see Fig. 10.25). In the inner, X-ray-dominated knots, the high-energy component overwhelms the low-energy one.

The second component can be attributed to either synchrotron radiation by a second population of high-energy electrons (or protons), or the beamed IC emission by the radio-emitting electrons. In the first case, the double synchrotron nature may arise from the presence of distinctively different acceleration processes (e.g. shock and turbulent acceleration). In the context of the origin of extragalactic cosmic-rays, it is interesting that a faster acceleration mechanism producing the second component is capable of accelerating cosmic-ray protons to energies $10^{16}\ \text{eV} \leq E_{p,\max} \leq 10^{19}\ \text{eV}$.

Problems

10.1. MRI: Derive the complete dispersion relation for the magnetorotational instability in a differentially rotating disk (see [49]). Discuss the various branches of the dispersion relation.

10.2. Ring Diffusion: Show that the solution

$$\Sigma(\tau, x) = \frac{m}{\pi R_0^2} \frac{1}{\tau x^{1/4}} \exp\left[-\frac{1+x^2}{\tau}\right] I_{1/4}(2x/\tau) \quad (10.178)$$

solves the diffusion equation (10.122). Study the asymptotic behavior of this solution in the limit $2x/\tau \ll 1$ and $2x/\tau \gg 1$. Derive the radial drift velocity for these two limits.

10.3. Relativistic Standard Disks: The Shakura–Sunyaev disk models presented in Sect. 10.4 are nonrelativistic and, hence, can only be viewed as a crude approximation to real black hole accretion disks near the horizon. The formal relativistic generalization of the steady-state Shakura–Sunyaev disk model to a relativistic accretion disk in a Kerr metric was made in 1974 by Kip Thorne and Don Page [320] (see also [342]). Formulate the relativistic angular momentum and energy conservation for geometrically thin disks and derive the relativistic dissipation function.

Show that for a nonrotating black hole the radiative efficiency is only $\eta \approx 0.06$, while a standard disk around a Kerr black hole with $a/M = 1$ has an efficiency of $\eta \approx 0.42$.

10.4. Radiative Transfer around Rotating Black Holes: Geometrically thin disks are best described in terms of cylindrical coordinates (t, ϕ, R, z) , such that $z = 0$ corresponds to the equatorial plane. The metric tensor of the Kerr geometry can now be expanded in powers of z/R .

- Derive the expressions for the metric tensor of the Kerr geometry up to second order (see also [342]).
- Show that the flow field U for matter in a Keplerian disk is given in leading order by

$$U^t = \frac{1}{\sqrt{\mathcal{B}}} \left(1 + \frac{a}{R} \sqrt{\frac{M}{R}} \right), \quad U^\phi = \frac{1}{\sqrt{\mathcal{B}}} \sqrt{\frac{M}{R^3}}, \quad (10.179)$$

where we use the two structure functions

$$\mathcal{A} = 1 - \frac{2M}{R} + \frac{a^2}{R^2} \quad (10.180)$$

$$\mathcal{B} = 1 - \frac{3M}{R} + \frac{2a}{R} \sqrt{\frac{M}{R}}. \quad (10.181)$$

With the velocity field $U = (U^t, U^\phi, 0, 0)$ one can now define a local rotating reference frame (LRFM), which can be extended to form a local tetrad \mathbf{e}_a with

$$\mathbf{e}_a = L^\beta_a B_\beta, \quad B_\alpha = \bar{L}^b_\alpha \mathbf{e}_b, \quad (10.182)$$

where $B_0 = U$ and B_i form a triad.

(iii) Determine the form of the transformation matrices L and \bar{L} .

The transport of radiation in the disk is described in terms of a Boltzmann equation for the photon distribution $f(x, p)$ depending on the coordinates x^μ and the photon four-momentum p^α (see Sect. 3.2). The phase-space distribution is related to the specific intensity I_ν by means of the relation $I_\nu = h\nu h^3 v^2 f(x, p)/c^2$. Since photons travel along null geodesics, except for interactions, the relativistic transport of photons follows from the geodesic spray

$$p^\alpha \frac{\partial f}{\partial x^\alpha} + \Gamma_{\alpha\beta}^i p^\alpha p^\beta \frac{\partial f}{\partial p^i} = Q, \quad (10.183)$$

where Q contains all possible interactions of photons with matter (emission, absorption and scattering). These calculations have to be done in the LRFM, since all interactions only have well known expressions in the local rest frame of the plasma. For this reason, we write the transfer equation in terms of the photon momenta as measured in the LRFM, $\bar{p}^\alpha = \bar{L}^\alpha_\beta p^\beta$,

$$L_\alpha^\alpha \bar{p}^\alpha \frac{\partial f}{\partial x^\alpha} + \omega_{ab}^i \bar{p}^a \bar{p}^b \frac{\partial f}{\partial \bar{p}^i} = \bar{Q}. \quad (10.184)$$

The connection coefficients ω_{ab}^i follow from the transformation

$$\omega_{ab}^i = \bar{L}_\varrho^i \left(L_a^\sigma \Gamma_{\sigma\gamma}^\varrho + \frac{\partial L_a^\varrho}{\partial x^\gamma} \right) L_b^\gamma. \quad (10.185)$$

The photon momentum is given by its frequency ν and direction \mathbf{n} , $\bar{p} = h\nu(1, \mathbf{n})$ with

$$\mathbf{n} = \left(\sqrt{1 - \mu^2} \sin \chi, \sqrt{1 - \mu^2} \cos \chi, \mu \right), \quad (10.186)$$

where μ is the cosine of the angle between the photon direction and z -axis, and χ is the corresponding azimuthal propagation angle on the unit sphere.

- (iv) Derive from this transformation an expression for the three-vector $m^i = \omega_{ab}^i \bar{p}^a \bar{p}^b$.
- (v) Derive an explicit expression for the stationary transfer equation (10.184) in order z/R .
- (vi) Discuss the diffusion approximation for geometrically thin disks.
- (vii) Find an expression for the operator \bar{Q} , when Bremsstrahlung and electron scattering are the dominant processes (see [13]).

10.5. GRMHD in Minkowski Space: Evaluate the GRMHD equations (10.167) in the limit of flat space. Compare your result with [251].

10.6. Characteristic Speeds of GRMHD: Consider the following set of variables, $\mathbf{V} = (U^\mu, b^\mu, P, s)$, where s is the specific entropy. Then GRMHD equations can be written as a quasilinear system of the form

$$\mathcal{A}_B^{\mu A} \nabla_\mu V^B = 0 \quad (10.187)$$

where A and B run from 0 to 9. Calculate the form of the matrices \mathcal{A}^μ . Let $\Phi(x^\mu) = 0$ define a characteristic hypersurface. Determine the characteristic matrix $\mathcal{A}^\mu \nabla_\mu$ and its determinant. Determine from this the characteristic waves (see also [33]).

11 Epilogue and Future Prospects

White dwarfs, neutron stars and black holes are fascinating objects in the Universe. In the last 20 years, great efforts have been undertaken to understand these compact objects in greater details. As we have outlined in this book, even the interior structure of white dwarfs is still somewhat uncertain. While the interior of black holes is of no astrophysical relevance, the internal structure of massive neutron stars is still under great debate. The internal structure of black holes is, however, of some importance for the understanding of Einstein's gravity. Inside the horizon, one should find a true vacuum with a vanishing energy-momentum tensor. From a modern point of view, vacuum has a structure – as we now know from the cosmological vacuum. There is some evidence that the Big Bang singularity will be avoided by quantum gravitational effects. Therefore, during the collapse of a stellar object towards a singular state we expect to run through the inverse of a Big Bang. In my opinion, the formation of black holes is by no means understood from a modern physical point of view.

The case for neutron stars is in a much better shape than black holes – except that neutron stars are probably not neutron stars, but either hybrid neutron stars or quark–neutron stars. About 2000 such objects appear in catalogs of radio pulsars and X-ray binary systems. For many of them, the masses have been measured quite accurately and found to cluster in a narrow range around 1.4 solar masses. This fact alone is not understood. In contrast to white dwarfs, which show a broad distribution in mass around a mean mass of 0.6 solar masses, this narrow peak distribution for neutron stars has to be explained by future supernovae theories.

White Dwarfs

A detailed record of the physical processes that operate during post-main-sequence evolution is contained in the internal chemical structure of white dwarfs. Global pulsations allow us to probe the stellar interior through asteroseismology, revealing the signatures of prior nuclear burning, mixing, and diffusion in these stars. The coolest white dwarfs in any stellar population can be used to constrain the age, but significant uncertainties remain in our understanding of the fundamental cooling physics. The hydrogen-atmosphere variable white dwarfs provide a unique opportunity to reduce these uncertainties empirically. Stellar crystallization occurs in the cores of these white dwarfs, releasing latent heat and delaying the gradual cooling of the star. The emphasis on white dwarf stars and cosmology arises from the most

recent advances in cosmological and galactic structure research in which white dwarf stars are playing a very prominent role. Examples are Type Ia supernovae (i.e. white dwarf supernovae), the origin and evolution of the Universe, the age of the galactic disk, cosmochronology using white dwarfs in globular clusters and galactic clusters, and the physics of accretion onto compact stars. In addition, GAIA will open up the possibility to test the mass–radius relation for white dwarfs in a much broader context.

When the white dwarf accumulates enough mass, about 1.4 solar masses, it will obliterate itself. This catastrophic event is a Type Ia supernova. Because a precise amount of mass is involved, every Type Ia supernova in the Universe shines at a unified, known luminosity.

Neutron Stars

In the 40 years that have elapsed since the discovery of pulsars, rapidly rotating highly magnetized neutron stars, the study of these fascinating objects has resulted in many applications in physics and astronomy. Striking examples include the confirmation of the existence of gravitational radiation, as predicted by general relativity, and the first detection of an extrasolar planetary. Pulsar astronomy remains an extremely active area of modern astrophysics and the next decade will undoubtedly continue to produce new results from currently known objects as well as new surprises.

Through an understanding of the Galactic population of radio pulsars, it is possible to predict the detection statistics of terrestrial gravitational wave detectors to nearby rapidly spinning neutron stars, as well as coalescing relativistic binaries at cosmic distances. Continued improvements in gravitational wave detector sensitivities should result in a number of interesting developments and contributions in this area. These developments and contributions might include the detection of presently known radio pulsars, as well as a population of coalescing binary systems which have not yet been detected as radio pulsars. The phenomenal timing stability of radio pulsars leads naturally to a large number of applications, including their use as laboratories for relativistic gravity and as natural detectors of gravitational radiation. Long-term timing experiments of the present sample of millisecond and binary pulsars currently underway appear to have tremendous potential in these areas and perhaps detect the gravitational wave background (if it exists) within the next decade.

These applications will benefit greatly from the continued discovery of new systems by the present generation of radio pulsar searches which continue to probe new areas of parameter space. It is clear that we are aware of only about 1% of the total active pulsar population in our Galaxy. It is therefore likely that we have not seen all of the pulsar zoo. More sensitive surveys are being planned both in the short term and in the longer term (the Square Kilometer Array). These should provide a far more complete census of the Galactic pulsar population. Possible discoveries in the future include:

- More dual-line binary pulsars, i.e. a double neutron star systems in which both components are observable as radio pulsars. The additional clock in such a binary system would be most valuable in further tests of strong-field gravity.

- A radio pulsar with a black-hole companion would undoubtedly also be a fantastic laboratory for studying gravity in the strong-field regime.
- A submillisecond pulsar. The original millisecond pulsar, B1937+21, rotating at 642 Hz is still the second most rapidly rotating neutron star known. Do kHz neutron stars exist? Searches now have sensitivity to such objects and a discovery of even one would constrain the equation of state of matter at high densities.
- A binary system in which the neutron star is in the process of transforming from an X-ray-emitting neutron star to a millisecond radio pulsar.

Relativistic binaries are tracers for the rich dynamical evolution of globular clusters. The populations of these objects are the result of an interplay between the gravitational dynamics of large N -body systems, the dynamics of mass transfer, the details of stellar evolution, and the effect of the gravitational field of the galaxy. The gravitational dynamics of globular clusters can enhance the population of short period binaries of main-sequence stars as well as inject compact objects such as white dwarfs and neutron stars into stellar binary systems. Once they are in such systems, the details of stellar evolution and mass transfer in close binary systems govern the likely end products of the dynamical interaction between the two stars. Furthermore, most models of the evolution of the core of a globular cluster rely on the gradual hardening and ejection of binary systems to delay the onset of core collapse. The hardening of binaries in the core of globular clusters will produce relativistic binaries, but it will also eventually eject these systems as they gain larger and larger recoil velocities in each subsequent encounter. The threshold for ejection from a globular cluster depends both upon the gravitational potential of the cluster itself and the gravitational potential of its environment generated by the Milky Way. As the globular cluster orbits the Milky Way, its local environment changes. Consequently, if other dynamical processes (such as gravothermal oscillations) do not dominate, the globular cluster's population of relativistic binaries may also reflect the past orbital history of the globular cluster.

Rotating relativistic stars are of fundamental interest in physics. Their bulk properties constrain the proposed equations of state for densities greater than nuclear density. Accreted matter in their gravitational fields undergoes high-frequency oscillations that could become a sensitive probe for general relativistic effects. Temporal changes in the rotational period of millisecond pulsars can also reveal a wealth of information about important physical processes inside the stars or of cosmological relevance. In addition, rotational instabilities can produce gravitational waves, the detection of which would initiate a new field of observational asteroseismology of relativistic stars.

Recent advances in numerical relativity have enabled the long-term dynamical evolution of rotating stars. Several interesting phenomena, such as dynamical instabilities, pulsation modes, and neutron star and black hole formation in rotating collapse have now been studied in full general relativity. The current studies are limited to relativistic polytropes, but new 3D simulations with realistic equations of state should be expected in the near future.

Collisions of black holes and neutron stars are interesting because of at least two reasons. Firstly, it is expected that they emit a large amount of energy as gravitational waves, which could be measured by new detectors. The form of those waves is expected to carry information about the internal structure of such systems. Secondly, collisions of such objects are the prime suspects of short gamma-ray bursts. The exact mechanism for the energy emission is unknown so far.

Black Holes

Black holes have been found so far in two mass intervals, either as stellar black holes with masses in the range of 2.5 to about 30 solar masses, or as supermassive black holes with masses in the range of millions to billions of solar masses. Our Galaxy is expected to harbor about 100,000 stellar black holes, only about two dozens have been identified as X-ray binary systems. None has been found so far as companion to a neutron star or white dwarf. In contrast to this, every massive galaxy harbors a supermassive black hole in its center. Since the observable Universe contains at least 100 billion massive galaxies, a tremendous number of potential black hole candidates will appear in future galaxy surveys. In fact, in previous surveys more than 100,000 black hole candidates have been found in quasars. Black holes are in fact no longer exotic objects, but are now considered to be real compact objects in the Universe. One interesting aspect of these objects is their spin state. Similar to neutron stars, black holes can be moderately rotating objects, or very rapidly rotating objects when accreting gas from their environment. The measurement of the spin of black holes is certainly one of the key issues for future research.

A black hole is a kind of ground state for a self-gravitating object. With the advent of the Kerr solution in 1963 and the postulate that black holes are behind the enormous energy output of quasars, a new era has been opened up in the physics of compact objects. For the past 20 years, astronomers have looked for stellar mass black holes by measuring binary orbits and supermassive black holes by analyzing rotation and random velocities of stars and gas near galactic centers. If the velocities are large enough, as, for example, in the Sombrero Galaxy, then they imply more mass than we see in stars. The only explanation which is in agreement with present-day physics is a black hole. About 50 have been found in nearby galaxies. Their masses are in the range expected for nuclear engines, and their numbers are consistent with predictions based on the energy output of quasars. In the next years, with new telescope techniques this number of black holes should considerably increase.

So far, dark clusters as an alternative to supermassive black holes have been ruled out in only two galaxies, NGC 4258 and our Milky Way. These two galaxies give us important proof that black holes exist. But both are special cases – NGC 4258 contains a disk of water masers that are observed with radio telescopes, and our Galactic center is so close that one can follow individual stellar orbits. Andromeda is now the first galaxy in which exotic alternatives to a black hole using Hubble can be excluded.

With an ever-increasing number of secure detections, supermassive black holes (SMBHs) have evolved, in the span of a few years, from exotic possibilities to

well-established components of galaxies. While it was understood since the 1960s that the energy sources of quasars must be gravitational, it was 30 years before the existence of SMBHs was firmly established, through measurements of the Keplerian rise in the rotation velocity of stars or gas at the very centers of galactic nuclei. It is now generally accepted that the formation and evolution of galaxies and SMBHs are tightly intertwined, from the early phases of protogalactic formation, through hierarchical build-up in CDM-like cosmogonies, to recent galaxy mergers.

On small scales, SMBHs are embedded in stellar cusps, parsec-scale regions where the stellar density increases approximately as a power law with distance from the SMBH into the smallest resolvable radii. Faint galaxies have steep nuclear density profiles, while bright galaxies typically have weaker cusps. Steep cusps form naturally as the growth of the SMBH pulls in stars. In small dense galaxies where the star-star relaxation time is shorter than 10^{10} yr, steep cusps may also form via collisional relaxation. Weak cusps may be remnants of strong cusps that were destroyed by binary SMBHs during galaxy mergers; in fact the structure and kinematics of galactic nuclei are now believed to be fossil relics of the merger process.

The proximity of our Galaxy's center presents a unique opportunity to study the environment of a supermassive black hole with much higher spatial resolution than can be brought to bear on any other galaxy. In 1995, this research has been initiated

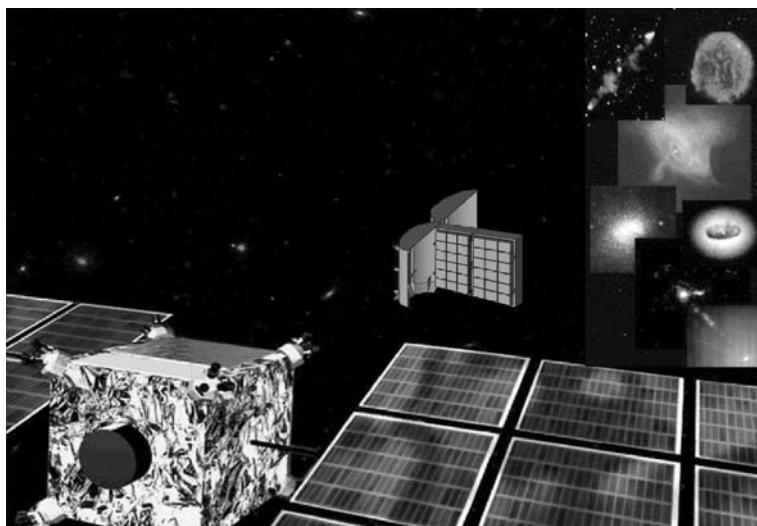


Fig. 11.1. Black hole explorers in the future. Black holes and neutron stars in close binary systems are promising sources of X-rays. XEUS will be a permanent space-borne X-ray observatory with a sensitivity comparable to the most advanced planned future facilities such as JWST, ALMA and Herschel. The mission is under study as envisaged by the Horizons 2000 Survey Committee. XEUS will be around 200 times more sensitive than XMM–Newton. Credit: image credit ESA

a diffraction-limited study on the Keck 10-meter telescope, of the Galaxy's central cluster.

One of the most intriguing puzzles to stem from a definitive case for a supermassive black hole at the center of the Milky Way is the origin of the apparently young stars in the central parsec; the existing gas in this region is far from being sufficiently dense for self-gravity to overcome the strong tidal forces from the black hole, making star formation an unlikely process unless past densities in the central parsec were far more extreme than at present. A wide range of possible solutions have been proposed to account for the apparently young stars in the vicinity of a supermassive black hole. One possibility is that the local gas density was much higher in the past, allowing gravitational collapse to occur in the presence of a strong tidal field. This could in principle occur through a pre-existing accretion disk, which has since dispersed or by the collision of infalling dense gas clouds although no such sufficiently dense gas clouds have been observed. Alternatively, stars might form at larger radii, where the tidal forces are much lower, and migrate inwards through dynamical friction. For this to work, the stars would have to belong to a stellar cluster so dense that it would have undergone core collapse, possibly forming an intermediate mass black hole. In a dramatically different approach, it has also been suggested that these stars are not truly young but are old stars that have been altered by their environment. At present, none of the proposed theories is altogether satisfactory, leaving both the He I emission-line stars and the Sgr A* cluster stars as paradoxes of apparent youth in the vicinity of a supermassive black hole.

GRMHD and Accretion

The description of many important areas of modern astronomy, such as high-energy astrophysics or gravitational wave astronomy, requires general relativity. High-energy radiation is often emitted by highly relativistic events in regions of strong gravitational fields near compact objects such as neutron stars or black holes. The production of relativistic radio jets in active galactic nuclei, explained by pure hydrodynamical effects as in the twin-exhaust model, by hydromagnetic centrifugal acceleration as in the Blandford–Payne mechanism, or by electromagnetic extraction of energy as in the Blandford–Znajek mechanism, involves an accretion disk around a rotating supermassive black hole. The discovery of kHz quasiperiodic oscillations in low-mass X-ray binaries extended the frequency range over which these oscillations occur into timescales associated with the relativistic, innermost regions of accretion disks. A relativistic description is also necessary in scenarios involving explosive collapse of very massive stars to a black hole (in the so-called collapsar and hypernova models), or during the last phases of the coalescence of neutron star binaries. These catastrophic events are believed to exist at the central engine of highly energetic gamma-ray bursts (GRBs). In addition, nonspherical gravitational collapse leading to black hole formation or to a supernova explosion, and neutron star binary coalescence are among the most promising sources of detectable gravitational radiation. Such astrophysical scenarios constitute one of the main targets for the new

generation of ground-based laser interferometers, just starting their gravitational wave search (LIGO, VIRGO, GEO600, TAMA).

A powerful way to improve our understanding of the above scenarios is through accurate, large-scale, three-dimensional numerical simulations. Nowadays, computational general relativistic astrophysics is an increasingly important field of research. In addition to the large amount of observational data gathered by high-energy X- and gamma-ray satellites such as Chandra, XMM–Newton, or INTEGRAL, and the new generation of gravitational wave detectors, the rapid increase in computing power through parallel supercomputers and the associated advance in software technologies is making possible large scale numerical simulations in the framework of general relativity. In the most general case, the equations governing the dynamics of relativistic astrophysical systems are an intricate, coupled system of time-dependent partial differential equations, comprising the (general) relativistic (magneto-)hydrodynamic (MHD) equations and the Einstein gravitational field equations. In many cases, the number of equations must be augmented to account for nonadiabatic processes, e.g. radiative transfer or sophisticated microphysics (realistic equations of state for nuclear matter, nuclear physics, magnetic fields, and so on). In this field of research we expect great progress in the next years.

A Astrophysical Constants and Symbols

Physical Constants

Quantity	Symbol	Value [SI]
Speed of light	c	$299\,792\,458 \text{ m s}^{-1}$
Newtonian gravitational constant	G	$6.6742(10) \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$
Planck constant	h	$6.6260693(10) \times 10^{-34} \text{ J s}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.05457168 \times 10^{-34} \text{ J s}$
Planck constant	$\hbar c$	$197.326968 \text{ MeV fm}$
Boltzmann constant	k_B	$1.380658 \times 10^{-23} \text{ J/K}$
Electron mass	m_e	$9.1093897 \times 10^{-31} \text{ kg}$
Electron charge	e	$1.60217733 \times 10^{-19} \text{ C}$
Proton mass	m_p	$1.67262158 \times 10^{-27} \text{ kg}$
Neutron mass	m_n	$1.6749286 \times 10^{-27} \text{ kg}$
Unified atomic mass unit	m_u	$1.6605402 \times 10^{-27} \text{ kg}$
Radiation constant	$a_{\text{SB}} = \pi^2 k_B^4 / 15 c^3 \hbar^3$	$7.56 \times 10^{-23} \text{ J m}^{-3} \text{ K}^{-4}$
Stefan–Boltzmann constant	$\sigma_{\text{SB}} = c a_{\text{SB}} / 4$	$5.6704 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$
Fine structure constant	$\alpha = e^2 / 4\pi\epsilon_0\hbar c$	$1/137.0359895$
Classical electron radius	$r_e = e^2 / 4\pi\epsilon_0 m_e c^2$	$2.81794092 \times 10^{-15} \text{ m}$
Thomson cross-section	$\sigma_T = 8\pi r_e^2 / 3$	$6.65246154 \times 10^{-29} \text{ m}^2$
Nuclear radius	$R_0 = 1.2 A^{1/3} \text{ fm}$	$1 \text{ fm} = 10^{-15} \text{ m}$
Nuclear saturation density	n_0	0.1620 fm^{-3}

Astronomical Quantities

Quantity	Symbol	Value [SI]
Astronomical unit	AU	$1.4959787066 \times 10^{11}$ m
Parsec	pc	$3.0856775807 \times 10^{16}$ m
Sidereal year	365.25636042 d	3.1558150×10^7 s
Solar mass	M_{\odot}	1.98892×10^{30} kg
Solar luminosity	L_{\odot}	3.846×10^{26} W
Solar radius (equatorial)	R_{\odot}	6.961×10^6 m
Schwarzschild radius of the Sun	$R_S = 2GM_{\odot}/c^2$	2.95325008 km
Solar system unit of time	$T_{\odot} = GM_{\odot}/c^3$	$4.92549047 \mu\text{s}$
Eddington luminosity	$L_{\text{Ed}} = 4\pi GMm_pc/\sigma_T$	$1.257 \times 10^{31} M/M_{\odot}$ W
Critical magnetic field	$B_{\text{crit}} = m_e^2 c^3 / \hbar e$	4.4×10^9 T
Chandrasekhar mass	$M_{\text{Ch}} = (5.87/\mu^2)M_{\odot}$	$1.457(2/\mu)^2 M_{\odot}$
Gravitational wave energy loss	$L_0 = c^5/G$	3.628×10^{52} W
Planck mass	$m_P = \sqrt{\hbar c/G}$	1.22090×10^{19} GeV/ c^2
Planck length	$L_P = \sqrt{\hbar G/c^3}$	1.61624×10^{-35} m
Planck time	$t_P = \sqrt{\hbar G/c^5}$	5.39121×10^{-44} s
Planck charge	$q_P = \sqrt{4\pi\epsilon_0\hbar c}$	1.87554×10^{-18} C
Planck current	$I_P = q_P/t_P$	3.4789×10^{25} A
Planck voltage	$V_P = \hbar/t_P q_P$	1.04295×10^{27} V
Planck impedance	$Z_P = V_P/I_P$	29.9792 Ω

List of Symbols

Physical Variable	Symbol	Typical Unit
Redshift factor	α	Dimensionless
Metric of three-space	γ_{ik}	Dimensionless
Shift vector field	β	
Exterior curvature	K_{ik}	Square inverse length
Observer (tetrad) field	e_a	Inverse length
One-form basis	Θ^a	Length
Specific angular momentum of black hole	a	Length
Angular momentum of black hole	a_*	Dimensionless
Lagrangian	\mathcal{L}	Energy
Lie derivative along X	L_X	Inverse length
Mass of black hole	M_H	Solar mass
Gravitational radius	M	Length
Covariant derivative along X	∇_X	Inverse length
Connection one-forms	ω_b^a	Inverse length
Curvature two-forms	Ω_b^a	Square inverse length
Grand canonical potential	\mathcal{Q}	Energy
Christoffel symbol	$\Gamma_{\alpha\beta}^\mu$	Inverse length
Ricci tensor	R_{ab}	Square inverse length
Einstein tensor	G_{ab}	Square inverse length
Energy–momentum tensor	T^{ab}	Energy density
Faraday tensor	F_{ab}	Electric field
Bardeen (Newtonian) potential	Φ	Dimensionless
Magnetic flux function	Ψ	Magnetic flux
Vector field on manifold	X	Inverse length
Periastron shift of binary orbit	$\dot{\omega}$	Degrees per revolution
Redshift and Doppler amplitude	γ_{RD}	Time
Shapiro range parameter	r	Time
Shapiro inclination parameter	s	Dimensionless
State vector of primitive variables	\mathbf{P}	
State vector of conserved variables	\mathbf{U}	
Flux vector of conserved variables	\mathbf{F}	
Lorentz factor	W	Dimensionless

Abbreviations and Acronyms

Symbol	Meaning
AGN	Active Galactic Nucleus
AMR	Adaptive Mesh Refinement
ASCA	Japanese X-Ray Satellite
BH	Black Hole
CFL	Courant–Friedrichs–Lewy condition
CHANDRA	CHANDRAsekhara X-ray Observatory (NASA)
CHOMBO	Block-Structured Adaptive Mesh Refinement Library
DE200	Development Ephemeris 200
EoS	Equation of State
ESA	European Space Agency
EUVN	EUropean VLbi Network
GAIA	Satellite named after a Greek Earth Goddess (ESA)
GP-B	Gravity Probe B
GRMHD	General Relativistic MagnetoHydrodynamics
HIPPARCOS	HIgh Precision PARallax Collecting Satellite
INTEGRAL	INTErnational Gamma-Ray Astrophysics Laboratory
JD	Julian Date
JWST	James Webb Space Telescope (NASA)
LAGEOS	LAser GEodetic Satellite
LMXB	Low-Mass X-Ray Binary System
MJD	Modified Julian Date
MPI	Message Passing Interface
NS	Neutron Star
PARAMESH	Parallel Adaptive Mesh Refinement
QSO	Quasistellar Object
QSR	Quasistellar Radio Source
ROSITA	ROentgen Survey with an Imaging Telescope Array
RXTE	Rossi X-ray Timing Explorer
SEP	Strong Equivalence Principle
SRMHD	Special Relativistic MagnetoHydrodynamics
VLBA	Very Long Baseline Array (USA)
VLBI	Very Long Baseline Interferometry
VLTI	Very Large Telescopes (ESO)
WD	White Dwarf
WEP	Weak Equivalence Principle
XEUS	X-ray Evolving Universe Spectroscopy Mission (ESA)
XMM–NEWTON	X-ray Multi-Mirror Satellite (ESA)
ZAMO	Zero Angular Momentum Observer

B SLy4 Equation of State for Neutron Star Matter

The equation of state (EoS) of dense neutron star matter is one of the mysteries of these objects. The EoS is a basic input for construction of neutron star models. Its knowledge is needed to calculate various properties of neutron stars. The EoS is predominantly determined by the nuclear (strong) interaction between elementary constituents of dense matter. Even in the neutron star crust, with density below normal nuclear density $\varrho_0 = 2.7 \times 10^{14} \text{ g cm}^{-3}$ (corresponding to baryon density $n_0 = 0.16 \text{ fm}^{-3}$), nuclear interactions are responsible for the properties of neutron rich nuclei, crucial for the crust EoS. The knowledge of these interactions is particularly important for the structure of the inner neutron star crust, where nuclei are immersed in a neutron gas, and even more so for the EoS of the liquid core. Nuclear interactions are actually responsible for a dramatic lifting of M_{\max} from $0.7 M_\odot$, obtained when interactions are switched off, to more realistic values of $1.4 M_\odot$ as measured in neutron star binary systems.

In the following tables, an equation of state of neutron star matter, describing both the neutron star crust and the liquid core, is given from the paper [136]. It is based on the effective nuclear interaction SLy of the Skyrme type, which is particularly suitable for the application to the calculation of the properties of very neutron rich matter.

Table B.1. Structure and composition of the inner neutron-star crust. For caption, see next table

n_b (fm $^{-3}$)	Z	A	X_n	R_p (fm)	R_n (fm)	R_{cell} (fm)	u (%)
1.2126 E-4	42.198	130.076	0.0000	5.451	5.915	63.503	0.063
1.6241 E-4	42.698	135.750	0.0000	5.518	6.016	58.440	0.084
1.9772 E-4	43.019	139.956	0.0000	5.565	6.089	55.287	0.102
2.0905 E-4	43.106	141.564	0.0000	5.578	6.111	54.470	0.107
2.2059 E-4	43.140	142.161	0.0247	5.585	6.122	54.032	0.110
2.3114 E-4	43.163	142.562	0.0513	5.590	6.128	53.745	0.113
2.6426 E-4	43.215	143.530	0.1299	5.601	6.145	53.020	0.118
3.0533 E-4	43.265	144.490	0.2107	5.612	6.162	52.312	0.123
3.5331 E-4	43.313	145.444	0.2853	5.623	6.179	51.617	0.129
4.0764 E-4	43.359	146.398	0.3512	5.634	6.195	50.937	0.135
4.6800 E-4	43.404	147.351	0.4082	5.645	6.212	50.269	0.142
5.3414 E-4	43.447	148.306	0.4573	5.656	6.228	49.615	0.148
6.0594 E-4	43.490	149.263	0.4994	5.667	6.245	48.974	0.155
7.6608 E-4	43.571	151.184	0.5669	5.690	6.278	47.736	0.169
1.0471 E-3	43.685	154.094	0.6384	5.725	6.328	45.972	0.193
1.2616 E-3	43.755	156.055	0.6727	5.748	6.362	44.847	0.211
1.6246 E-3	43.851	159.030	0.7111	5.784	6.413	43.245	0.239
2.0384 E-3	43.935	162.051	0.7389	5.821	6.465	41.732	0.271
2.6726 E-3	44.030	166.150	0.7652	5.871	6.535	39.835	0.320
3.4064 E-3	44.101	170.333	0.7836	5.923	6.606	38.068	0.377
4.4746 E-3	44.155	175.678	0.7994	5.989	6.698	36.012	0.460
5.7260 E-3	44.164	181.144	0.8099	6.059	6.792	34.122	0.560
7.4963 E-3	44.108	187.838	0.8179	6.146	6.908	32.030	0.706
9.9795 E-3	43.939	195.775	0.8231	6.253	7.048	29.806	0.923
1.2513 E-2	43.691	202.614	0.8250	6.350	7.171	28.060	1.159
1.6547 E-2	43.198	211.641	0.8249	6.488	7.341	25.932	1.566
2.1405 E-2	42.506	220.400	0.8222	6.637	7.516	24.000	2.115
2.4157 E-2	42.089	224.660	0.8200	6.718	7.606	23.106	2.458
2.7894 E-2	41.507	229.922	0.8164	6.825	7.721	22.046	2.967
3.1941 E-2	40.876	235.253	0.8116	6.942	7.840	21.053	3.585
3.6264 E-2	40.219	240.924	0.8055	7.072	7.967	20.128	4.337
3.9888 E-2	39.699	245.999	0.7994	7.187	8.077	19.433	5.058
4.4578 E-2	39.094	253.566	0.7900	7.352	8.231	18.630	6.146
4.8425 E-2	38.686	261.185	0.7806	7.505	8.372	18.038	7.202
5.2327 E-2	38.393	270.963	0.7693	7.685	8.538	17.499	8.470
5.6264 E-2	38.281	283.993	0.7553	7.900	8.737	17.014	10.011
6.0219 E-2	38.458	302.074	0.7381	8.167	8.987	16.598	11.914
6.4183 E-2	39.116	328.489	0.7163	8.513	9.315	16.271	14.323
6.7163 E-2	40.154	357.685	0.6958	8.853	9.642	16.107	16.606
7.0154 E-2	42.051	401.652	0.6699	9.312	10.088	16.058	19.501
7.3174 E-2	45.719	476.253	0.6354	9.990	10.753	16.213	23.393
7.5226 E-2	50.492	566.654	0.6038	10.701	11.456	16.557	26.996
7.5959 E-2	53.162	615.840	0.5898	11.051	11.803	16.772	28.603

Table B.2. Previous table: Structure and composition of the inner neutron-star crust (ground state) calculated within the compressible liquid drop model with SLy effective nucleon-nucleon interaction. X_n is the fraction of nucleons in the neutron gas outside nuclei. Upper part with $X_n = 0$ corresponds to a shell of the outer crust, just above the neutron drip surface in the neutron-star interior, and calculated within the same model. R_p and R_n are the equivalent proton and neutron radii. Wigner-Seitz cell radius and fraction of volume occupied by nuclear matter (equal to that occupied by protons) are denoted by R_{cell} and u , respectively. This table: Equation of state of the inner crust. First line corresponds to the neutron drip point, as calculated within the COMPRESSIBLE LIQUID DROP MODEL. Last line corresponds to the bottom edge of the crust

n_b (fm $^{-3}$)	ρ (g cm $^{-3}$)	P (erg cm $^{-3}$)	Γ	n_b (fm $^{-3}$)	ρ (g cm $^{-3}$)	P (erg cm $^{-3}$)	Γ
2.0905 E-4	3.4951 E11	6.2150 E29	1.177	9.9795 E-3	1.6774 E13	3.0720 E31	1.342
2.2059 E-4	3.6883 E11	6.4304 E29	0.527	1.2513 E-2	2.1042 E13	4.1574 E31	1.332
2.3114 E-4	3.8650 E11	6.5813 E29	0.476	1.6547 E-2	2.7844 E13	6.0234 E31	1.322
2.6426 E-4	4.4199 E11	6.9945 E29	0.447	2.1405 E-2	3.6043 E13	8.4613 E31	1.320
3.0533 E-4	5.1080 E11	7.4685 E29	0.466	2.4157 E-2	4.0688 E13	9.9286 E31	1.325
3.5331 E-4	5.9119 E11	8.0149 E29	0.504	2.7894 E-2	4.7001 E13	1.2023 E32	1.338
4.0764 E-4	6.8224 E11	8.6443 E29	0.554	3.1941 E-2	5.3843 E13	1.4430 E32	1.358
4.6800 E-4	7.8339 E11	9.3667 E29	0.610	3.6264 E-2	6.1153 E13	1.7175 E32	1.387
5.3414 E-4	8.9426 E11	1.0191 E30	0.668	3.9888 E-2	6.7284 E13	1.9626 E32	1.416
6.0594 E-4	1.0146 E12	1.1128 E30	0.726	4.4578 E-2	7.5224 E13	2.3024 E32	1.458
7.6608 E-4	1.2831 E12	1.3370 E30	0.840	4.8425 E-2	8.1738 E13	2.6018 E32	1.496
1.0471 E-3	1.7543 E12	1.7792 E30	0.987	5.2327 E-2	8.8350 E13	2.9261 E32	1.536
1.2616 E-3	2.1141 E12	2.1547 E30	1.067	5.6264 E-2	9.5022 E13	3.2756 E32	1.576
1.6246 E-3	2.7232 E12	2.8565 E30	1.160	6.0219 E-2	1.0173 E14	3.6505 E32	1.615
2.0384 E-3	3.4178 E12	3.7461 E30	1.227	6.4183 E-2	1.0845 E14	4.0509 E32	1.650
2.6726 E-3	4.4827 E12	5.2679 E30	1.286	6.7163 E-2	1.1351 E14	4.3681 E32	1.672
3.4064 E-3	5.7153 E12	7.2304 E30	1.322	7.0154 E-2	1.1859 E14	4.6998 E32	1.686
4.4746 E-3	7.5106 E12	1.0405 E31	1.344	7.3174 E-2	1.2372 E14	5.0462 E32	1.685
5.7260 E-3	9.6148 E12	1.4513 E31	1.353	7.5226 E-2	1.2720 E14	5.2856 E32	1.662
7.4963 E-3	1.2593 E13	2.0894 E31	1.351	7.5959 E-2	1.2845 E14	5.3739 E32	1.644

Table B.3. Top: Composition of the liquid core. Fractions of particles are defined as $x_j = n_j/n_b$. Neutron fraction can be calculated using $x_n = 1 - x_p$. Bottom: Equation of state of the liquid neutron star core

n_b (fm $^{-3}$)	x_p (%)	x_e (%)	x_μ (%)	n_b (fm $^{-3}$)	x_p (%)	x_e (%)	x_μ (%)
0.0771	3.516	3.516	0.000	0.490	7.516	4.960	2.556
0.0800	3.592	3.592	0.000	0.520	7.587	4.954	2.634
0.0850	3.717	3.717	0.000	0.550	7.660	4.952	2.708
0.0900	3.833	3.833	0.000	0.580	7.736	4.955	2.781
0.1000	4.046	4.046	0.000	0.610	7.818	4.964	2.854
0.1100	4.233	4.233	0.000	0.640	7.907	4.979	2.927
0.1200	4.403	4.398	0.005	0.670	8.003	5.001	3.002
0.1300	4.622	4.521	0.101	0.700	8.109	5.030	3.079
0.1600	5.270	4.760	0.510	0.750	8.309	5.094	3.215
0.1900	5.791	4.896	0.895	0.800	8.539	5.178	3.361
0.2200	6.192	4.973	1.219	0.850	8.803	5.284	3.519
0.2500	6.499	5.014	1.485	0.900	9.102	5.410	3.692
0.2800	6.736	5.031	1.705	0.950	9.437	5.557	3.880
0.3100	6.920	5.034	1.887	1.000	9.808	5.726	4.083
0.3400	7.066	5.026	2.040	1.100	10.663	6.124	4.539
0.3700	7.185	5.014	2.170	1.200	11.661	6.602	5.060
0.4000	7.283	4.999	2.283	1.300	12.794	7.151	5.643
0.4300	7.368	4.984	2.383	1.400	14.043	7.762	6.281
0.4600	7.444	4.971	2.473	1.500	15.389	8.424	6.965

n_b (fm $^{-3}$)	ϱ (g cm $^{-3}$)	P (erg cm $^{-3}$)	Γ	n_b (fm $^{-3}$)	ϱ (g cm $^{-3}$)	P (erg cm $^{-3}$)	Γ
0.0771	1.3038 E14	5.3739 E32	2.159	0.4900	8.8509 E14	1.0315 E35	2.953
0.0800	1.3531 E14	5.8260 E32	2.217	0.5200	9.4695 E14	1.2289 E35	2.943
0.0850	1.4381 E14	6.6828 E32	2.309	0.5500	1.0102 E15	1.4491 E35	2.933
0.0900	1.5232 E14	7.6443 E32	2.394	0.5800	1.0748 E15	1.6930 E35	2.924
0.1000	1.6935 E14	9.9146 E32	2.539	0.6100	1.1408 E15	1.9616 E35	2.916
0.1100	1.8641 E14	1.2701 E33	2.655	0.6400	1.2085 E15	2.2559 E35	2.908
0.1200	2.0350 E14	1.6063 E33	2.708	0.6700	1.2777 E15	2.5769 E35	2.900
0.1300	2.2063 E14	1.9971 E33	2.746	0.7000	1.3486 E15	2.9255 E35	2.893
0.1600	2.7223 E14	3.5927 E33	2.905	0.7500	1.4706 E15	3.5702 E35	2.881
0.1900	3.2424 E14	5.9667 E33	2.990	0.8000	1.5977 E15	4.2981 E35	2.869
0.2200	3.7675 E14	9.2766 E33	3.025	0.8500	1.7302 E15	5.1129 E35	2.858
0.2500	4.2983 E14	1.3668 E34	3.035	0.9000	1.8683 E15	6.0183 E35	2.847
0.2800	4.8358 E14	1.9277 E34	3.032	0.9500	2.0123 E15	7.0176 E35	2.836
0.3100	5.3808 E14	2.6235 E34	3.023	1.0000	2.1624 E15	8.1139 E35	2.824
0.3400	5.9340 E14	3.4670 E34	3.012	1.1000	2.4820 E15	1.0609 E36	2.801
0.3700	6.4963 E14	4.4702 E34	2.999	1.2000	2.8289 E15	1.3524 E36	2.778
0.4000	7.0684 E14	5.6451 E34	2.987	1.3000	3.2048 E15	1.6876 E36	2.754
0.4300	7.6510 E14	7.0033 E34	2.975	1.4000	3.6113 E15	2.0679 E36	2.731
0.4600	8.2450 E14	8.5561 E34	2.964	1.5000	4.0498 E15	2.4947 E36	2.708

C 3+1 Split of Spacetime Curvature

In this appendix we derive the Gauss equation and the Codazzi–Mainardi equations for the 3+1 decomposition of the Riemann curvature.

C.1 Gauss Decomposition

In analogy to the decomposition of the connection form discussed in Sect. 2.8.3 we split the curvature two-form given by the second structure equation

$$\Omega^a_b = d\omega^a_b + \omega^a_c \wedge \omega^c_b. \quad (\text{C.1})$$

The Gauss decomposition follows from the spatial part of the second structure equation

$$\begin{aligned} \Omega^i_j &= d\omega^i_j + \omega^i_k \wedge \omega^k_j + \omega^i_0 \wedge \omega^0_j \\ &= d\left[\bar{\omega}^i_j(e_k)\bar{\Theta}^k + H^i_j\Theta^0\right] + (\nabla_i \ln \alpha \Theta^0 - K^i_k \Theta^k) \wedge (\nabla_j \ln \alpha \Theta^0 - K_{jm} \Theta^m) \\ &\quad + (\bar{\omega}^i_k + H^i_k \Theta^0) \wedge (\bar{\omega}^k_j + H^k_j \Theta^0) \\ &= \bar{d}\bar{\omega}^i_j + \bar{\omega}^i_k \wedge \bar{\omega}^k_j \\ &\quad + d(H^i_j \Theta^0) - \nabla_i \ln \alpha K_{jm} \Theta^0 \wedge \Theta^m + \nabla_j \ln \alpha K^i_m \Theta^0 \wedge \Theta^m \\ &\quad + K^i_m K_{jk} \Theta^m \wedge \Theta^k + H^k_j \omega^i_k \wedge \Theta^0 + H^i_k \Theta^0 \wedge \omega^k_j \\ &= \bar{\Omega}^i_j + K^i_k K_{jm} \Theta^k \wedge \Theta^m \\ &\quad - [\nabla_i \ln \alpha K_{jm} - \nabla_j \ln \alpha K^i_m] \Theta^0 \wedge \Theta^m \\ &\quad + [\nabla_k H^i_j + \bar{\omega}^i_k H^k_j - H^i_k \bar{\omega}^k_j] \wedge \Theta^0 \\ &= \bar{\Omega}^i_j + K^i_k K_{jm} \Theta^k \wedge \Theta^m \\ &\quad - [\nabla_i(\ln \alpha) K_{jm} - \nabla_j(\ln \alpha) K^i_m + D_m H^i_j] \Theta^0 \wedge \Theta^m. \end{aligned} \quad (\text{C.2})$$

This shows that we find for the curvature on the three-surface Σ

$$\begin{aligned}
\Omega^i_j|_{\Sigma} &= [d\omega^i_j + \omega^i_s \wedge \omega^s_j + \omega^i_0 \wedge \omega^0_j]|_{\Sigma} \\
&= d\bar{\omega}^i_j + \bar{\omega}^i_s \wedge \bar{\omega}^s_j + (-\eta^{im} K_{ms} \bar{\Theta}^s) \wedge (-K_{jt} \bar{\Theta}^t) \\
&= \bar{\Omega}^i_j + (-K^i_s \bar{\Theta}^s) \wedge (-K_{jt} \bar{\Theta}^t) \\
&= \bar{\Omega}^i_j + K^i_s K_{jt} \bar{\Theta}^s \wedge \bar{\Theta}^t. \tag{C.3}
\end{aligned}$$

This Gauss equation expresses the 3D curvature tensor in terms of the projection of the 4D curvature, with extrinsic curvature corrections. In fact, the expressions for the Ricci tensor given in the next section show that the second part of Ω^i_j is not needed for the calculation of the Ricci tensor.

C.2 Codazzi–Mainardi Equations

The Codazzi–Mainardi equation follows from

$$\begin{aligned}
\Omega_i^0 &= d\omega_i^0 + \omega_k^0 \wedge \omega_i^k \\
&= d[\nabla_i \ln \alpha \Theta^0 - K_{ij} \Theta^j] + [(\nabla_k \ln \alpha) \Theta^0 - K_{kj} \Theta^j] \wedge \omega_i^k \\
&= -\frac{1}{\alpha} d\alpha, i \wedge \Theta^0 - \nabla_k \ln \alpha \omega_i^k \wedge \Theta^0 \\
&\quad - dK_{ij} \wedge \Theta^j - K_{ij} d\Theta^j + K_{kj} \omega_i^k \wedge \Theta^j \\
&= \frac{1}{\alpha} [d\alpha, i - \alpha, k \omega_i^k] \wedge \Theta^0 - dK_{ij} \wedge \Theta^j + K_{ik} \omega_a^k \wedge \Theta^a + K_{kj} \omega_i^k \wedge \Theta^j \\
&= \frac{1}{\alpha} D(\alpha, i) \wedge \Theta^0 - dK_{ij} \wedge \Theta^j + K_{ij} \omega_0^k \wedge \Theta^0 \\
&\quad + K_{ik} \omega_m^k \wedge \Theta^m + K_{kj} \omega_i^k \wedge \Theta^j \\
&= \frac{1}{\alpha} D(\alpha, i) \wedge \Theta^0 \\
&\quad - [dK_{ij} - K_{ik} \omega_j^k - K_{kj} \omega_i^k] \wedge \Theta^j - K_{ik} K_m^k \Theta^m \wedge \Theta^0. \tag{C.4}
\end{aligned}$$

This decomposition provides then the projection of the 4D curvature Ω_i^0 known as the Codazzi–Mainardi equation

$$\Omega_i^0|_{\Sigma} = -\bar{D}K_{ij} \wedge \bar{\Theta}^j, \tag{C.5}$$

where $\bar{D}K_{ij} = (D_s K_{ij}) \bar{\Theta}^s$.

For the calculation of the Ricci tensor we only need the normal projections

$$\Omega_0^i(e_j, e_0) = \Omega_i^0(e_j, e_0). \tag{C.6}$$

From the above we derive for this

$$\Omega_i^0(e_j, e_0) = \frac{1}{\alpha} D_j(\alpha, i) + dK_{ij}(e_0) - K_{im} \omega_j^m(e_0) - K_{kj} \omega_i^k(e_0) - \mathbf{K}_{ij}^2. \tag{C.7}$$

Remember that

$$dK_{ij}(e_0) = \frac{1}{\alpha} (\partial_t - i_\beta \cdot d) K_{ij}. \quad (\text{C.8})$$

Now we consider the term, following from equation (2.368)

$$\begin{aligned} (K_\omega)_{ij} &= K_{im}\omega_j^m(e_0) + K_{jm}\omega_i^m(e_0) \\ &= \frac{1}{\alpha} K_i^m [\beta_{[m|j]} - c_{[m|j]} - \bar{\omega}_{mj}(\beta)] \\ &\quad + \frac{1}{\alpha} K_j^m [\beta_{[m|i]} - c_{[mi]} - \bar{\omega}_{mi}(\beta)] \\ &= -\frac{1}{\alpha} [K_{im}\bar{\omega}_j^m(\beta) + K_{jm}\bar{\omega}_i^m(\beta)] \\ &\quad + \frac{1}{2\alpha} [K_i^m \beta_{m|j} - K_i^m \beta_{j|m} + K_j^m \beta_{m|i} - K_j^m \beta_{i|m} \\ &\quad - K_i^m c_{mj} + K_i^m c_{jm} - K_j^m c_{mi} + K_j^m c_{im}] \\ &= -\frac{1}{\alpha} [K_{im}\bar{\omega}_j^m(\beta) + K_{jm}\bar{\omega}_i^m(\beta)] \\ &\quad + \frac{1}{2\alpha} [-K_i^m (\beta_{m|j} + \beta_{j|m} - (c_{mj} + c_{jm})) - 2K_i^m c_{mj} \\ &\quad - K_j^m (\beta_{m|i} + \beta_{i|m} - (c_{mi} + c_{im})) - 2K_j^m c_{im} \\ &\quad + 2K_i^m \beta_{m|j} + 2K_j^m \beta_{m|i}] . \end{aligned} \quad (\text{C.9})$$

In this expression, all other terms including c_{im} cancel out. Using the definition of the extrinsic curvature K_{ij} , equation (2.367), we find

$$\begin{aligned} (K_\omega)_{ij} &= -K_{im}K_j^m - K_j^m K_{mi} + \frac{1}{\alpha} K_{im}\beta_{|j}^m + \frac{1}{\alpha} K_{jm}\beta_{|i}^m \\ &\quad - \frac{1}{\alpha} [K_{im}\bar{\omega}_j^m(\beta) + K_{jm}\bar{\omega}_i^m(\beta)] \\ &= -K_{im}K_j^m - K_j^m K_{mi} + \frac{1}{\alpha} K_{im}\beta_{,j}^m + \frac{1}{\alpha} K_{jm}\beta_{,i}^m . \end{aligned} \quad (\text{C.10})$$

Together with the expression for the Lie derivative of the extrinsic curvature

$$\mathcal{L}_\beta K_{ij} = \beta^m K_{ij,m} + K_{im}\beta_{,j}^m + K_{mj}\beta_{,i}^m \quad (\text{C.11})$$

we found for the curvature component

$$\begin{aligned} \mathcal{Q}_i^0(e_j, e_0) &= \frac{1}{\alpha} D_j(\alpha_{,i}) + \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) K_{ij} \\ &\quad + K_{im}K_j^m + K_{jm}K_i^m - \mathbf{K}_{ij}^2 \\ &= \frac{1}{\alpha} D_j(\alpha_{,i}) + \frac{1}{\alpha} (\partial_t - \mathcal{L}_\beta) K_{ij} + K_{jm}K_i^m . \end{aligned} \quad (\text{C.12})$$

D 3+1 Split of Rotating Neutron Star Geometry

D.1 The 3+1 Split of the Connection

We will apply Cartan's methods to calculate the curvature tensor for rotating space-times with respect to Bardeen observers. In a first step we calculate the exterior derivatives for the fundamental one-forms (7.16) ($A, B = 2, 3$)

$$\begin{aligned} d\Theta^A &= \sum_B \exp(\mu_A) \mu_{A,B} dx^B \wedge dx^A + \exp(\mu_A) \mu_{A,\phi} d\phi \wedge dx^B \\ &= \exp(-\mu_B) \mu_{A,B} \Theta^B \wedge \Theta^A. \end{aligned} \quad (\text{D.1})$$

Similarly, we find

$$d\Theta^1 = \sum_A \exp(-\mu_A) \psi_{,A} \Theta^A \wedge \Theta^1 - \sum_A \exp(\psi - \nu - \mu_A) \omega_{,A} \Theta^A \wedge \Theta^0, \quad (\text{D.2})$$

as well as

$$d\Theta^0 = \sum_A \exp(-\mu_A) \nu_{,A} \Theta^A \wedge \Theta^0. \quad (\text{D.3})$$

Here we used the inversion

$$d\phi = \exp(-\psi) \Theta^1 + \exp(\psi - \nu) \omega \Theta^0. \quad (\text{D.4})$$

Comparing this with Cartan's first structure equations

$$d\Theta^0 = - \sum_A \omega_A^0 \wedge \Theta^A - \omega_1^0 \wedge \Theta^1 \quad (\text{D.5})$$

$$d\Theta^1 = - \sum_A \omega_A^1 \wedge \Theta^A - \omega_0^1 \wedge \Theta^0 \quad (\text{D.6})$$

and

$$d\Theta^A = - \sum_B \omega_B^A \wedge \Theta^B - \omega_1^A \wedge \Theta^1 - \omega_0^A \wedge \Theta^0, \quad (\text{D.7})$$

we conclude for axisymmetric connections

$$\omega_A^1 = -\omega_A^A = \exp(-\mu_A)\psi_{,A}\Theta^1 - \exp(-\psi)\mu_{,A}\Theta^A \quad (\text{D.8})$$

$$\omega_B^A = -\omega_A^B = \exp(-\mu_B)\mu_{A,B}\Theta^A - \exp(-\mu_A)\mu_{B,A}\Theta^B. \quad (\text{D.9})$$

The following ansatz solves the structure equations for the six connection forms of the Lorentz connection for axisymmetric and stationary spacetimes ($\omega_i^0 = \omega_0^i$, $\omega_j^i = -\omega_i^j$)

$$\omega_1^0 = -\frac{1}{2}\exp(\psi - v - \mu_2)\omega_{,2}\Theta^2 - \frac{1}{2}\exp(\psi - v - \mu_3)\omega_{,3}\Theta^3 \quad (\text{D.10})$$

$$\omega_2^0 = \exp(-\mu_2)v_{,2}\Theta^0 - \frac{1}{2}\exp(\psi - v - \mu_2)\omega_{,2}\Theta^1 \quad (\text{D.11})$$

$$\omega_3^0 = \exp(-\mu_3)v_{,3}\Theta^0 - \frac{1}{2}\exp(\psi - v - \mu_3)\omega_{,3}\Theta^1 \quad (\text{D.12})$$

$$\omega_2^1 = \exp(-\mu_2)\psi_{,2}\Theta^1 + \frac{1}{2}\exp(\psi - v - \mu_2)\omega_{,2}\Theta^0 \quad (\text{D.13})$$

$$\omega_3^1 = \exp(-\mu_3)\psi_{,3}\Theta^1 + \frac{1}{2}\exp(\psi - v - \mu_3)\omega_{,3}\Theta^0 \quad (\text{D.14})$$

$$\omega_3^2 = \exp(-\mu_3)\mu_{2,3}\Theta^2 - \exp(-\mu_2)\mu_{3,2}\Theta^3. \quad (\text{D.15})$$

These relations can be contracted in a way which shows the features of the general decomposition found in Sect. 2.8

$$\omega_A^0 = \nabla_A \ln(\alpha)\Theta^0 - K_{A1}\Theta^1 \quad (\text{D.16})$$

$$\omega_A^1 = (\nabla_A\psi)\Theta^1 + \frac{1}{2\alpha}(R\nabla_A\omega)\Theta^0 = \bar{\omega}_A^1 + \frac{1}{2\alpha}(R\nabla_A\omega)\Theta^0 \quad (\text{D.17})$$

$$\omega_3^2 = (\nabla_3\mu_2)\Theta^2 - (\nabla_2\mu_3)\Theta^3 = \bar{\omega}_3^2, \quad (\text{D.18})$$

where K_{ij} is the extrinsic curvature (remember that $c_{ij} \equiv 0$ in a stationary spacetime), given in orthonormal basis,

$$K_{\hat{i}\hat{j}} = \frac{1}{2\alpha}(\beta_{i;j} + \beta_{j;i}). \quad (\text{D.19})$$

With the definition of the covariant derivative for $\beta = (-\omega \exp \Psi, 0, 0)$

$$\beta_{i;j} = e_j(\beta) - \omega_i^m(e_j)\beta_m \quad (\text{D.20})$$

we obtain the following form for the extrinsic curvature

$$K_{\hat{i}\hat{j}} = -\frac{R}{2\alpha} \begin{pmatrix} 0 & \nabla_2\omega & \nabla_3\omega \\ \nabla_2\omega & 0 & 0 \\ \nabla_3\omega & 0 & 0 \end{pmatrix}, \quad (\text{D.21})$$

which shows that $\text{Tr}(\mathbf{K}) = 0$. $\nabla_A\alpha = e_A^\mu\partial_\mu\alpha$ denotes the derivative along the meridional vector field \mathbf{e}_A and $R = \exp(\psi)$ is the cylindrical radius.

The expressions for the connection one-forms are just a special case of the general 3+1 split derived in Sect. 2.8

$$\omega_i^0 = (\nabla_i \ln \alpha) \Theta^0 - K_{ij} \Theta^j \quad (\text{D.22})$$

$$\omega_j^i = \bar{\omega}_j^i + H_j^i \Theta^0, \quad (\text{D.23})$$

where the matrix \mathbf{H} is antisymmetric (for stationary spacetimes)

$$H_{ij} = \frac{1}{\alpha} \beta_{[i|j]}. \quad (\text{D.24})$$

β is often called the vector potential of the gravitomagnetic field and the antisymmetric matrix \mathbf{H} defines then the gravitomagnetic field itself, quite in analogy to the magnetic field in electrodynamics. All components of the connection therefore have a quite clear physical or geometrical interpretation.

D.2 The Curvature of Time Slices

For the calculation of the Ricci tensor of the hypersurface we need the curvature of the meridional plane

$$\bar{\Omega}_3^2 = d\bar{\omega}_3^2 + \bar{\omega}_1^2 \wedge \bar{\omega}_3^1 \quad (\text{D.25})$$

and the curvature of the other 3D directions

$$\bar{\Omega}_A^1 = d\bar{\omega}_A^1 + \bar{\omega}_B^1 \wedge \bar{\omega}_A^B. \quad (\text{D.26})$$

For this purpose we define two poloidal vectors ($A = 2, 3$)

$$Q_A \equiv \exp(-\mu_A) \omega_{,A}, \quad \Psi_A \equiv \exp(-\mu_A) \psi_{,A} = \nabla_A \psi. \quad (\text{D.27})$$

In terms of these quantities we can write for any function $F(x^2, x^3)$

$$\begin{aligned} d(F\bar{\Theta}^1) &= \sum_A \exp(-\psi - \mu_A) (F \exp \psi)_{,A} \bar{\Theta}^A \wedge \bar{\Theta}^1 \\ &= \sum_A \frac{1}{R} \nabla_A [RF] \bar{\Theta}^A \wedge \bar{\Theta}^1 \end{aligned} \quad (\text{D.28})$$

$$\begin{aligned} d(F\bar{\Theta}^A) &= \sum_B \exp(-\mu_A - \mu_B) (\exp \mu_A F)_{,B} \bar{\Theta}^B \wedge \bar{\Theta}^A \\ &= \sum_B \exp(-\mu_A) \nabla_B [\exp(\mu_A) F] \bar{\Theta}^B \wedge \bar{\Theta}^A. \end{aligned} \quad (\text{D.29})$$

So we need the exterior derivatives of ω_2^1 , ω_3^1 and of ω_3^2 . Using this rule in conjunction with the above connection form we obtain

$$\begin{aligned} d\bar{\omega}_2^1 + \bar{\omega}_3^1 \wedge \bar{\omega}_2^3 &= \frac{1}{R} \nabla_A [R\Psi_2] \bar{\Theta}^A \wedge \bar{\Theta}^1 \\ &\quad - \Psi_3 \nabla_3 \mu_2 \bar{\Theta}^1 \wedge \bar{\Theta}^2 + \Psi_3 \nabla_2 \mu_3 \bar{\Theta}^1 \wedge \bar{\Theta}^3 \end{aligned} \quad (\text{D.30})$$

$$\begin{aligned} d\bar{\omega}_3^1 + \bar{\omega}_2^1 \wedge \bar{\omega}_3^2 &= \frac{1}{R} \nabla_A [R\Psi_3] \bar{\Theta}^A \wedge \bar{\Theta}^1 \\ &\quad + \Psi_2 \nabla_3 \mu_2 \bar{\Theta}^1 \wedge \bar{\Theta}^2 - \Psi_2 \nabla_2 \mu_3 \bar{\Theta}^1 \wedge \bar{\Theta}^3 \end{aligned} \quad (\text{D.31})$$

$$\begin{aligned} d\bar{\omega}_3^2 + \bar{\omega}_1^2 \wedge \bar{\omega}_3^1 &= - [\exp(-\mu_2) \nabla_3 [\exp(\mu_2) \nabla_3 \mu_2] \\ &\quad + \exp(-\mu_3) \nabla_2 [\exp(\mu_3) \nabla_2 \mu_3]] \bar{\Theta}^2 \wedge \bar{\Theta}^3. \end{aligned} \quad (\text{D.32})$$

With these expressions the curvature of the hypersurfaces can be written in closed form

$$\begin{aligned} \bar{\Omega}_2^1 = & - \left[\frac{1}{R} \nabla_2 [R\Psi_2] + \Psi_3 [\nabla_3 \mu_2] \right] \bar{\Theta}^1 \wedge \bar{\Theta}^2 \\ & + \left[-\frac{1}{R} \nabla_3 [R\Psi_2] + \Psi_3 [\nabla_2 \mu_3] \right] \bar{\Theta}^1 \wedge \bar{\Theta}^3 \end{aligned} \quad (\text{D.33})$$

$$\begin{aligned} \bar{\Omega}_3^1 = & - \left[\frac{1}{R} \nabla_3 [R\Psi_3] + \Psi_2 [\nabla_2 \mu_3] \right] \bar{\Theta}^1 \wedge \bar{\Theta}^3 \\ & + \left[-\frac{1}{R} \nabla_2 [R\Psi_3] + \Psi_2 [\nabla_3 \mu_2] \right] \bar{\Theta}^1 \wedge \bar{\Theta}^2 \end{aligned} \quad (\text{D.34})$$

$$\begin{aligned} \bar{\Omega}_3^2 = & - [\exp(-\mu_2) \nabla_3 [\exp(\mu_2) \nabla_3 \mu_2] \\ & + \exp(-\mu_3) \nabla_2 [\exp(\mu_3) \nabla_2 \mu_3]] \bar{\Theta}^2 \wedge \bar{\Theta}^3. \end{aligned} \quad (\text{D.35})$$

The curvature tensor of a three-surface has nine independent components and satisfies in our case $R_{1213} = R_{1312}$ and $R_{1223} = 0 = R_{1323}$.

With these expression we can calculate the six components of the Ricci tensor of the hypersurface, $\bar{R}_{ij} = \bar{\Omega}_i^m (e_m, e_j)$,

$$\begin{aligned} \bar{R}_{11} &= \bar{\Omega}_2^2(e_2, e_1) + \bar{\Omega}_3^1(e_3, e_1) = \bar{\Omega}_2^1(e_1, e_2) + \bar{\Omega}_3^1(e_1, e_3) = \\ &= -\frac{1}{R} \nabla_2 [R\Psi_2] - \Psi_3 [\nabla_3 \mu_2] - \frac{1}{R} \nabla_3 [R\Psi_3] - \Psi_2 [\nabla_2 \mu_3] \end{aligned} \quad (\text{D.36})$$

$$\begin{aligned} \bar{R}_{22} &= \bar{\Omega}_2^1(e_1, e_2) + \bar{\Omega}_3^2(e_3, e_2) \\ &= -\frac{1}{R} \nabla_2 [R\Psi_2] - \Psi_3 [\nabla_3 \mu_2] - \exp(-\mu_2) \nabla_3 [\exp(\mu_2) \nabla_3 \mu_2] \\ &\quad - \exp(-\mu_3) \nabla_2 [\exp(\mu_3) \nabla_2 \mu_3] \end{aligned} \quad (\text{D.37})$$

$$\begin{aligned} \bar{R}_{33} &= \bar{\Omega}_3^1(e_1, e_3) + \bar{\Omega}_3^2(e_2, e_3) \\ &= -\frac{1}{R} \nabla_3 [R\Psi_3] - \Psi_2 [\nabla_2 \mu_3] - \exp(-\mu_2) \nabla_3 [\exp(\mu_2) \nabla_3 \mu_2] \\ &\quad - \exp(-\mu_3) \nabla_2 [\exp(\mu_3) \nabla_2 \mu_3] \end{aligned} \quad (\text{D.38})$$

$$\bar{R}_{12} = \bar{\Omega}_2^1(e_2, e_2) + \bar{\Omega}_3^1(e_3, e_2) = 0 \quad (\text{D.39})$$

$$\bar{R}_{13} = \bar{\Omega}_1^2(e_2, e_3) + \bar{\Omega}_1^3(e_3, e_3) = 0 \quad (\text{D.40})$$

$$\bar{R}_{23} = \bar{\Omega}_2^1(e_1, e_3) + \bar{\Omega}_2^3(e_3, e_3) = -\frac{1}{R}\nabla_3[R\Psi_2] + \Psi_3[\nabla_2\mu_3]. \quad (\text{D.41})$$

By summation we get the Ricci scalar on the hypersurface

$$\begin{aligned} \bar{R} &= \bar{R}_{11} + \bar{R}_{22} + \bar{R}_{33} \\ &= -2 \left[\frac{1}{R}\nabla_2[R\Psi_2] + \Psi_3(\nabla_3\mu_2) + \frac{1}{R}\nabla_3[R\Psi_3] + \Psi_2(\nabla_2\mu_3) \right] \\ &\quad - 2\exp(-\mu_2)\nabla_3[\exp(\mu_2)\nabla_3\mu_2] - 2\exp(-\mu_3)\nabla_2[\exp(\mu_3)\nabla_2\mu_3] \\ &= -2 \left[\frac{1}{R}\nabla_2[R\Psi_2] + \Psi_3(\nabla_3\mu_2) + \frac{1}{R}\nabla_3[R\Psi_3] + \Psi_2(\nabla_2\mu_3) \right] \\ &\quad - 2\Delta(\mu_2, \mu_3), \end{aligned} \quad (\text{D.42})$$

where we have defined the second-order elliptic operator

$$\begin{aligned} \Delta(\mu_2, \mu_3) &= \exp(-\mu_2)\nabla_3[\exp(\mu_2)(\nabla_3\mu_2)] \\ &\quad + \exp(-\mu_3)\nabla_2[\exp(\mu_3)(\nabla_2\mu_3)]. \end{aligned} \quad (\text{D.43})$$

E Equations of GRMHD

E.1 Electromagnetic Fields

A complete description of the electromagnetic field is provided by the Faraday tensor $F^{\mu\nu}$, which is related to the electric and magnetic field, E^μ and B^μ , measured by an observer with four-velocity O

$$F^{\mu\nu} = O^\mu E^\nu - O^\nu E^\mu + O_\varrho \eta^{\varrho\mu\nu\sigma} B_\sigma. \quad (\text{E.1})$$

Both, electric and magnetic fields are orthogonal to O and are recovered from the Faraday tensor by means of the following relations

$$E^\mu = F^{\mu\nu} O_\nu \quad (\text{E.2})$$

and

$$B^\mu = \frac{1}{2} \eta^{\mu\nu\varrho\sigma} O_\nu F_{\varrho\sigma} = O_\nu * F^{\nu\mu}. \quad (\text{E.3})$$

The dual of the electromagnetic tensor is defined as

$$*F^{\mu\nu} = \frac{1}{2} \eta^{\mu\nu\varrho\sigma} F_{\varrho\sigma}, \quad (\text{E.4})$$

or expressed as

$$*F^{\mu\nu} = O^\mu B^\nu - O^\nu B^\mu + \eta^{\mu\nu\varrho\sigma} O_\varrho E_\sigma. \quad (\text{E.5})$$

$\eta^{\mu\nu\varrho\sigma} = [\mu\nu\varrho\sigma]/\sqrt{-g}$ is the total antisymmetric tensor related to the volume element (see Sect. 2.3).

We now decompose the Faraday tensor into electric and magnetic components measured by Eulerian observers \mathbf{n} by means of

$$E^\mu = F^{\mu\nu} n_\nu, \quad B^\mu = -* F^{\mu\nu} n_\nu. \quad (\text{E.6})$$

Both fields are purely spatial, $E^\mu n_\mu = 0 = B^\mu n_\mu$. This is equivalent to decompose the Faraday tensor into

$$F^{\mu\nu} = n^\mu E^\nu - n^\nu E^\mu + n_\varrho \eta^{\varrho\mu\nu\sigma} B_\sigma.$$

(E.7)

This electromagnetic part simplifies if we adopt the ideal MHD approximation: the electric field as measured in the plasma frame vanishes due to the high conductivity of the plasma. In SRMHD this is the famous relation $\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0$. The covariant expression for this condition is (Ohm's law)

$$U^\mu F_{\mu\nu} = 0. \quad (\text{E.8})$$

But even in GR, we may still define magnetic fields as measured in plasma frame

$$b^\mu = -\frac{1}{2} \eta^{\mu\nu\rho\sigma} U_\nu F_{\rho\sigma}. \quad (\text{E.9})$$

In the case of ideal MHD, this relation can easily be inverted to give

$$F^{\mu\nu} = U_\rho \eta^{\rho\mu\nu\sigma} b_\sigma. \quad (\text{E.10})$$

Taking the dual, we obtain

$$*F^{\mu\nu} = b^\mu U^\nu - b^\nu U^\mu. \quad (\text{E.11})$$

The magnetic field b^μ only lives in three-space, since $U_\mu b^\mu = 0$. Inserting this into the expression (10.160) yields the energy-momentum tensor of the electromagnetic part in terms of the comoving magnetic field

$$T_{(\text{ED})}^{\mu\nu} = \frac{b^2}{4\pi} U^\mu U^\nu + \frac{b^2}{8\pi} g^{\mu\nu} - \frac{1}{4\pi} b^\mu b^\nu. \quad (\text{E.12})$$

This expression is very similar to the classical EM tensor except for the contribution to the energy density. In summary, we have found the stress-energy tensor for a plasma

$$T^{\mu\nu} = \left(\varrho_0 + \epsilon + P + \frac{b^2}{4\pi} \right) U^\mu U^\nu + \left(P + \frac{b^2}{8\pi} \right) g^{\mu\nu} - \frac{1}{4\pi} b^\mu b^\nu. \quad (\text{E.13})$$

This is the stress-energy tensor which we need for the equations of motion.

We may find the relations between magnetic fields in the plasma frame and the observer's frame by defining a projection operator $P_{\mu\nu} = g_{\mu\nu} + U_\mu U_\nu$. Since b^μ is orthogonal to U^μ , we find $P_\nu^\mu b^\nu = b^\mu$. It follows therefore from the definition of B^μ that

$$P_\nu^\mu B^\nu = P_\nu^\mu n_\varrho (b^\varrho U^\nu - b^\nu U^\varrho) = -n_\varrho U^\varrho b^\mu. \quad (\text{E.14})$$

Hence we have

$$b^\mu = -\frac{P_\nu^\mu B^\nu}{n_\nu U^\nu}. \quad (\text{E.15})$$

We can now evaluate the time and spatial components

$$b^t = U_i B^i / \alpha = \frac{W(\mathbf{v} \cdot \mathbf{B})}{\alpha} \quad (\text{E.16})$$

$$b^i = \frac{B^i / \alpha + b^t U^i}{U^t} = \frac{B^i + W^2(\mathbf{v} \cdot \mathbf{B}) v^i}{W}, \quad (\text{E.17})$$

where $U^t = W/\alpha$. Finally, the modulus of the plasma magnetic field can be written as

$$b^2 = \frac{B^2 + \alpha^2(b^t)^2}{W^2} = \frac{B^2 + W^2(\mathbf{v} \cdot \mathbf{B})^2}{W^2}, \quad (\text{E.18})$$

where $B^2 = B_i B^i$.

Maxwell's equations follow from the homogeneous equations

$$\nabla_\nu * F^{\mu\nu} = 0 = \frac{1}{\sqrt{-g}} \partial_\nu [\sqrt{-g} * F^{\mu\nu}], \quad (\text{E.19})$$

where $\sqrt{-g} = \alpha\sqrt{\gamma}$. The time component gives the divergence condition

$$\frac{1}{\sqrt{\gamma}} \partial_i [\sqrt{\gamma} B^i] = 0.$$

(E.20)

The spatial components give the induction equation in conservative form

$$\frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} B^i] + \frac{1}{\sqrt{-g}} \partial_j [\sqrt{-g} (U^j b^i - U^i b^j)] = 0. \quad (\text{E.21})$$

On the other hand we also find

$$U^j b^i - U^i b^j = V^j B^i - V^i B^j, \quad (\text{E.22})$$

where $V^i = v^i - \beta^i/\alpha$. The induction equation can therefore be written in the form

$$\frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} B^i] + \frac{1}{\sqrt{-g}} \partial_j [\sqrt{-g} (V^j B^i - V^i B^j)] = 0.$$

(E.23)

This form of the induction equation is equivalent to the conservative formulation of the Newtonian MHD.

E.2 Conservative Formulation of GRMHD

Similar to the approach chosen to model pure hydrodynamical flows in Sect. 3.1, we shortly discuss the time evolution of magnetohydrodynamic fields based on a conservative schemes. Baryon number conservation gives

$$\frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} D] + \frac{1}{\sqrt{-g}} \partial_j [\sqrt{-g} D V^j] = 0, \quad (\text{E.24})$$

where $D = \varrho_0 \alpha U^t = \varrho_0 W$ is a relativistic mass density. Similar to the hydro case, we introduce the momentum fluxes measured by Eulerian observers

$$S_i = -n_\mu T_i^\mu = \alpha T_i^t = [\varrho_0 h + b^2/4\pi] W^2 v_i - \alpha b^t b_i/4\pi, \quad (\text{E.25})$$

as well as the total energy density

$$\tau = n_\mu n_\nu T^{\mu\nu} - D = \alpha^2 T^{tt} - D = \varrho h_* W^2 - P_T - \alpha^2 (b^t)^2 - D, \quad (\text{E.26})$$

where $P_T = P + b^2/8\pi$ is the total pressure in the plasma and $h_* = h + b^2/4\pi\varrho_0$ the total enthalpy. The system is completed by means of an equation of state in the form of $P = (\Gamma - 1)\varrho_0 e$.

The spatial components of the energy-momentum conservation provide momentum conservation

$$\frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} S_i] + \frac{1}{\sqrt{-g}} \partial_j [\sqrt{-g} T_i^j] = T^{\mu\nu} \left(\frac{\partial g_{vi}}{\partial x^\mu} - \Gamma_{v\mu}^\sigma g_{\sigma i} \right), \quad (\text{E.27})$$

and the time-component gives the energy equation

$$\begin{aligned} \frac{1}{\sqrt{-g}} \partial_t [\sqrt{\gamma} \tau] + \frac{1}{\sqrt{-g}} \partial_j [\sqrt{-g} (\alpha T^{tj} - D V^j)] \\ = \alpha \left(T^{ut} \frac{\partial \log \alpha}{\partial x^\mu} - T^{\mu\nu} \Gamma_{v\mu}^t \right). \end{aligned} \quad (\text{E.28})$$

The GRMHD equations have therefore the form of a hyperbolic system, similar to (3.32),

$$\frac{1}{\sqrt{-g}} \left(\frac{\partial [\sqrt{\gamma} \mathbf{U}]}{\partial t} + \frac{\partial [\sqrt{-g} \mathbf{F}^i]}{\partial x^i} \right) = \mathcal{S}, \quad (\text{E.29})$$

which are obtained by combining the plasma equations with the induction equation (E.23). The state vector of GRMHD now consists of eight variables

$$\mathbf{U} = (D, S_i, \tau, B^i)^T, \quad (\text{E.30})$$

explicitly given by the vector in the state space

$$\mathbf{U} = \begin{pmatrix} D \\ S_1 \\ S_2 \\ S_3 \\ \tau \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} = \begin{pmatrix} \varrho_0 W \\ (\varrho_0 h + b^2/4\pi) W^2 v_1 - \alpha b^t b_1/4\pi \\ (\varrho_0 h + b^2/4\pi) W^2 v_2 - \alpha b^t b_2/4\pi \\ (\varrho_0 h + b^2/4\pi) W^2 v_3 - \alpha b^t b_3/4\pi \\ (\varrho_0 h + b^2/4\pi) W^2 - P_T - \alpha^2 (b^t)^2 / 4\pi - D \\ B^1 \\ B^2 \\ B^3 \end{pmatrix}. \quad (\text{E.31})$$

The corresponding fluxes \mathbf{F} are now given by

$$\mathbf{F}^i = \begin{pmatrix} DV^i \\ S_1 V^i - b_1 B^i / 4\pi W + P_T \delta_1^i \\ S_2 V^i - b_2 B^i / 4\pi W + P_T \delta_2^i \\ S_3 V^i - b_3 B^i / 4\pi W + P_T \delta_3^i \\ \tau V^i + P_T v^i - \alpha b^t B^i / W \\ B^1 V^i - B^i V^1 \\ B^2 V^i - B^i V^2 \\ B^3 V^i - B^i V^3 \end{pmatrix}, \quad (\text{E.32})$$

where $V^i \equiv v^i - \beta^i / \alpha$. The energy-momentum tensor in the sources \mathcal{S} now includes both parts, plasma and electromagnetic fields

$$\mathcal{S} = \begin{pmatrix} 0 \\ T^{\mu\nu} \partial_\mu g_{v1} - \Gamma_{v\mu}^\varrho g_{\varrho 1} \\ T^{\mu\nu} \partial_\mu g_{v2} - \Gamma_{v\mu}^\varrho g_{\varrho 2} \\ T^{\mu\nu} \partial_\mu g_{v3} - \Gamma_{v\mu}^\varrho g_{\varrho 3} \\ \alpha (T^{\mu t} \partial_\mu \alpha - T^{\mu\nu} \Gamma_{v\mu}^t) \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (\text{E.33})$$

E.3 Numerical Schemes

Recovery of Primitive Variables

Conservative MHD schemes require methods to transform between conserved variables \mathbf{U} and primitive variables \mathbf{P} . The time integration of GRMHD determines the three-momenta

$$S_i = (\varrho_0 h + b^2 / 4\pi) W^2 v_i - \alpha b^t b_i / 4\pi, \quad (\text{E.34})$$

the mass-density D and the energy τ . The associated four-momentum vector defined as

$$P_\mu = -n_v T_\mu^v = \alpha T_\mu^t \quad (\text{E.35})$$

has then the following form

$$P_\mu = W(\varrho_0 h + b^2 / 4\pi) U_\mu - (P + b^2 / 8\pi) n_\mu - \alpha b^t b_\mu / 4\pi. \quad (\text{E.36})$$

It is useful to remember the two relations

$$b^2 = \frac{1}{W^2} (\mathbf{B}^2 + (U_\mu B^\mu)^2), \quad n_v B^v = -U_\mu B^\mu. \quad (\text{E.37})$$

Noble et al. [311] discuss the mathematical properties of the inverse transformation and present six numerical methods for performing the inversion. Comparisons between the methods are made using a survey over phase space, a two-dimensional explosion problem, and a general relativistic MHD accretion disk simulation.

In the first method, we solve two algebraic equations simultaneously for $H = W^2 h \varrho_0$ and \mathbf{v}^2 . The momentum vector can be written, using $\mathbf{B} \rightarrow \mathbf{B}/\sqrt{4\pi}$, and the relation (E.42) in the following form

$$\mathbf{S} = (H + \mathbf{B}^2) \mathbf{v} - \frac{(\mathbf{S} \cdot \mathbf{B})}{H} \mathbf{B} \quad (\text{E.38})$$

and the energy as

$$\tau = \frac{\mathbf{B}^2}{2}(1 + \mathbf{v}^2) + \frac{\mathbf{S} \cdot \mathbf{B}}{2H} + H - D - P(e, \varrho_0). \quad (\text{E.39})$$

The first equation can be solved to get \mathbf{v}^2 as an explicit function of H

$$\mathbf{v}^2(H) = \frac{\mathbf{S}^2 H^2 + (\mathbf{S} \cdot \mathbf{B})^2 (\mathbf{B}^2 + 2H)}{(\mathbf{B}^2 + H)^2 H^2}. \quad (\text{E.40})$$

The energy equation provides then a second relation if we adopt a simple EoS

$$\tau = \frac{\mathbf{B}^2}{2}(1 + \mathbf{v}^2) + \frac{\mathbf{S} \cdot \mathbf{B}}{2H} + H - D - \left(\frac{\Gamma - 1}{\Gamma} [(1 - \mathbf{v}^2)H - \varrho_0] \right). \quad (\text{E.41})$$

The final step is to find \mathbf{v} by using \mathbf{S} . Starting with the expressions for b^t , b^i and b^2 in the definition of \mathbf{S} one finds

$$\boxed{\mathbf{S} = (H + \mathbf{B}^2) \mathbf{v} - (\mathbf{v} \cdot \mathbf{B}) \mathbf{B}.} \quad (\text{E.42})$$

This shows that the relativistic momentum flow \mathbf{S} has, besides kinematic factors, two relativistic corrections, the magnetic energy density $\mathbf{B}^2/4\pi$ and a Poynting contribution $\propto \mathbf{v} \cdot \mathbf{B}$. Since $\mathbf{v} \cdot \mathbf{B} = (\mathbf{S} \cdot \mathbf{B})/H$, we can use this to solve for the velocity field in terms of conserved variables

$$\boxed{\mathbf{v} = \frac{1}{H + \mathbf{B}^2} \left[\mathbf{S} + \frac{(\mathbf{S} \cdot \mathbf{B})}{H} \mathbf{B} \right].} \quad (\text{E.43})$$

Koide et al. [231] have proposed an alternative procedure to solve a combined system for the variables $x = W - 1$ and $y = W(\mathbf{v} \cdot \mathbf{B})$ which is given as

$$x(x + 2) \left[\Gamma R x^2 + (2\Gamma R - d)x + \Gamma R - d + e + \frac{\Gamma}{2} y^2 \right] \quad (\text{E.44})$$

$$= (\Gamma x^2 + 2\Gamma x + 1)^2 [f^2(x + 1)^2 + 2\sigma y + 2\sigma xy + g^2 y^2]$$

$$\begin{aligned} & \left[\Gamma(R - g^2)x^2 + (2\Gamma R - 2\Gamma g^2 - d)x + \Gamma R - d + e - g^2 + \frac{\Gamma}{2} y^2 \right] y \\ &= \sigma(x + 1)(\Gamma x^2 + 2\Gamma x + 1), \end{aligned} \quad (\text{E.45})$$

where $R = D + \tau$, $d = (\Gamma - 1)D$, $e = (1 - \Gamma/2)\mathbf{B}^2/4\pi$, and $\sigma = \mathbf{B} \cdot \mathbf{S}$. These algebraic equations are solved at each grid point using a two-variable Newton–Raphson iteration method. The primitive variables are then reconstructed from x , y , D , \mathbf{S} , τ , and \mathbf{B} with the following expressions

$$W = 1 + x \quad (E.46)$$

$$P = \frac{(\Gamma - 1)[\tau - xD - (2 - 1/W^2)\mathbf{B}^2/8\pi + (y/W)^2/2]}{Wx(x + 2) + 1} \quad (E.47)$$

$$\mathbf{v} = \frac{\mathbf{S} + (y/W)\mathbf{B}}{D + [\tau + P + \mathbf{B}^2/2W^2 + (y/W)^2/2]}. \quad (E.48)$$

On Numerical Implementations

There are many possible ways to numerically integrate the GRMHD equations. As in the Newtonian case, nonconservative schemes enjoyed wide use in the astrophysical community (e.g. ZEUS3D and NIRVANA2). They permit the integration of the internal energy density ϵ rather than the total energy equation. This is advantageous in regions of a plasma flow where the internal energy is small compared to the total energy, which is in fact the common situation in nonrelativistic astrophysics. De Villiers and Hawley [133] have developed a nonconservative scheme of GRMHD following a ZEUS-like approach. Modern approaches to solve GRMHD are however based on the above conservative formulation. This guarantees a true momentum and energy conservation.

Since we update \mathbf{U} rather than \mathbf{P} , we must solve at the end of each timestep for $\mathbf{P}(\mathbf{U})$. This can be done in various ways. The simplest approach is to use Newton–Raphson routines with the value of \mathbf{P} given by the previous time-step as an initial guess. Here, only five equations need to be solved, since B^i are analytically given. The Newton–Raphson procedure requires an expensive evaluation of the Jacobian $\partial\mathbf{U}/\partial\mathbf{P}$ and is in general limited in accuracy, i.e. it is a source of numerical noise. The evaluation of $\mathbf{P}(\mathbf{U})$ is at the heart of each numerical procedure for solving SRMHD or GRMHD. This procedure must be robust and CPU friendly.

A further important step is the evaluation of the fluxes \mathbf{F} . Gammie et al. [166] use a MUSCL type scheme with HLL fluxes (Harten et al. [194]). The fluxes are defined at zone faces. A slope-limited linear extrapolation from the zone center gives the values \mathbf{P}_R and \mathbf{P}_L for the primitive variables at the right and left sides of each zone interface. From \mathbf{P}_R and \mathbf{P}_L one calculates the maximum right- and leftgoing wave speeds and the fluxes $\mathbf{F}_R = \mathbf{F}(\mathbf{P}_R)$ and $\mathbf{F}_L = \mathbf{F}(\mathbf{P}_L)$. In the PPM reconstruction scheme, a quartic polynomial interpolation is used to obtain the primitive variables to the left and right of the grid cell interface. The relativistic version of the PPM algorithm can be found in Marti and Müller [271].

The exact solution of the Riemann problem in special relativistic magnetohydrodynamics (SRMHD) is discussed in [172]. Both initial states leading to a set of only three waves analogous to the ones in relativistic hydrodynamics, as well as generic initial states leading to the full set of seven MHD waves are considered. Because of

its generality, the solution presented could serve as an important test for numerical codes solving the MHD equations in relativistic regimes.

Time-Stepping Procedure

To advance time-steps, the higher order algorithms discussed in Sect. 3.1 can be applied. They are not repeated here.

Constrained Transport

Shock-capturing schemes do not guarantee $\nabla \cdot \mathbf{B} = 0$ for all time-steps. Some constrained transport schemes are needed to maintain $\nabla \cdot \mathbf{B} = 0$. Procedures of this type are discussed by Toth [398]. The flux-interpolated constrained transport (flux-CT) scheme introduced by Toth [398] is quite favorable for coding. In this algorithm, the numerical flux of the induction equation computed at each point is replaced with a linear combination of the numerical fluxes computed at each point and neighboring points. This procedure does not require a staggered mesh.

F Solutions

Problems in Chapter 2

2.1 The exterior derivative of $\omega = *A$ is an n -form, given by

$$\begin{aligned} (d\omega)_{ai_1 \dots i_{n-1}} &= (d * A)_{ai_1 \dots i_{n-1}} \\ &= n \nabla_{[a} (A^b \eta_{bi_1 \dots i_{n-1}]}) \\ &= n \eta_{b[i_1 \dots i_{n-1}] \nabla_a} A^b . \end{aligned} \quad (\text{F.1})$$

Since

$$d\omega = f \eta = *f , \quad (\text{F.2})$$

with its dual

$$f = (-1)^s * *f = (-1)^s * d\omega , \quad (\text{F.3})$$

we obtain in our case

$$\begin{aligned} *d\omega &= *d * A \\ &= \frac{1}{n!} \eta^{ai_1 \dots i_{n-1}} (\eta_{b[i_1 \dots i_{n-1}] \nabla_a} A^b) \\ &= \frac{1}{(n-1)!} (-1)^s (n-1)! \delta_b^a \nabla_a A^b \\ &= (-1)^s \nabla_a A^a . \end{aligned} \quad (\text{F.4})$$

From the definition of the Levi-Civita tensor we obtain

$$d\omega = (\nabla_a A^a) \sqrt{|g|} d^n x . \quad (\text{F.5})$$

2.3 See [2].

2.4 For this, see classical textbooks on general relativity; see also Sect. 6.4.3.

Problems in Chapter 3

3.1 For the solution of this problem we use the covariant expression (3.32) and calculate the Christoffel symbols for the various coordinate systems. In cylindrical

coordinates (r, ϕ, z) , the equations of special relativistic hydrodynamics (3.11) are given by

$$\frac{\partial D}{\partial t} + \frac{1}{r} \frac{\partial(rDv_r)}{\partial r} + \frac{1}{r} \frac{\partial(Dv_\phi)}{\partial \phi} + \frac{\partial(Dv_z)}{\partial z} = 0 \quad (\text{F.6})$$

$$\frac{\partial S_r}{\partial t} + \frac{1}{r} \frac{\partial[r(S_r v_r + P)]}{\partial r} + \frac{1}{r} \frac{\partial(S_r v_\phi)}{\partial \phi} + \frac{\partial(S_r v_z)}{\partial z} = \frac{P}{r} + \frac{\varrho_0 h W^2 v_\phi^2}{r} \quad (\text{F.7})$$

$$\frac{\partial S_\phi}{\partial t} + \frac{1}{r} \frac{\partial(rS_\phi v_r)}{\partial r} + \frac{1}{r} \frac{\partial(S_\phi v_\phi + P)}{\partial \phi} + \frac{\partial(S_\phi v_z)}{\partial z} = -\frac{\varrho_0 h W^2 v_r v_\phi}{r} \quad (\text{F.8})$$

$$\frac{\partial S_z}{\partial t} + \frac{1}{r} \frac{\partial(rS_z v_r)}{\partial r} + \frac{1}{r} \frac{\partial(S_z v_\phi)}{\partial \phi} + \frac{\partial(S_z v_z + P)}{\partial z} = 0 \quad (\text{F.9})$$

$$\begin{aligned} & \frac{\partial \tau}{\partial t} + \frac{1}{r} \frac{\partial[r(S_r - Dv_r)]}{\partial r} \\ & + \frac{1}{r} \frac{\partial(S_\phi - Dv_\phi)}{\partial \phi} + \frac{\partial(S_z - Dv_z)}{\partial z} = 0. \end{aligned} \quad (\text{F.10})$$

The discretized equation for cylindrical coordinates is given as

$$\begin{aligned} \frac{d\mathbf{U}_{i,j,k}}{dt} = & -\frac{r_{i+1/2}\mathbf{F}_{i+1/2,j,k}^r - r_{i-1/2}\mathbf{F}_{i-1/2,j,k}^r}{r_i \Delta r} \\ & -\frac{\mathbf{F}_{i,j+1/2,k}^\phi - \mathbf{F}_{i,j-1/2,k}^\phi}{r_i \Delta \phi} \\ & -\frac{\mathbf{F}_{i,j,k+1/2}^z - \mathbf{F}_{i,j,k-1/2}^z}{\Delta z} + S_{i,j,k}, \end{aligned} \quad (\text{F.11})$$

3.2 In spherical coordinates (r, θ, ϕ) , the equations of special relativistic hydrodynamics (3.11) are given by

$$\frac{\partial D}{\partial t} + \frac{1}{r^2} \frac{\partial(r^2 Dv_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta Dv_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial(Dv_\phi)}{\partial \phi} = 0 \quad (\text{F.12})$$

$$\begin{aligned} \frac{\partial S_r}{\partial t} + \frac{1}{r^2} \frac{\partial[r^2(S_r v_r + P)]}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta S_r v_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial(S_r v_\phi)}{\partial \phi} \\ = \frac{2P}{r} + \frac{\varrho_0 h W^2(v_\theta^2 + v_\phi^2)}{r} \end{aligned} \quad (\text{F.13})$$

$$\begin{aligned} \frac{\partial S_\theta}{\partial t} + \frac{1}{r^2} \frac{\partial(r^2 S_\theta v_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial[\sin \theta(S_\theta v_\theta + P)]}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial(S_\theta v_\phi)}{\partial \phi} \\ = \frac{P \cot \theta}{r} - \frac{\varrho_0 h W^2(v_\phi^2 \cot \theta - v_r v_\theta)}{r} \end{aligned} \quad (\text{F.14})$$

$$\begin{aligned} \frac{\partial S_\phi}{\partial t} + \frac{1}{r^2} \frac{\partial(r^2 S_\phi v_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(\sin \theta S_\phi v_\theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial(S_\phi v_\phi + P)}{\partial \phi} \\ = -\frac{\varrho_0 h W^2 v_\phi (v_r + v_\theta \cot \theta)}{r} \end{aligned} \quad (\text{F.15})$$

$$\begin{aligned} \frac{\partial \tau}{\partial t} + \frac{1}{r^2} \frac{\partial[r^2(S_r - Dv_r)]}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial[\sin \theta(S_\theta - Dv_\theta)]}{\partial \theta} \\ + \frac{1}{r \sin \theta} \frac{\partial(S_\phi - Dv_\phi)}{\partial \phi} = 0. \end{aligned} \quad (\text{F.16})$$

The discretized equation for spherical coordinates is given as

$$\begin{aligned} \frac{d\mathbf{U}_{i,j,k}}{dt} = & -\frac{r_{i+1/2}^2 \mathbf{F}_{i+1/2,j,k}^r - r_{i-1/2}^2 \mathbf{F}_{i-1/2,j,k}^r}{r_i^2 \Delta r} \\ & - \frac{\sin \theta_{j+1/2} \mathbf{F}_{i,j+1/2,k}^\theta - \sin \theta_{j-1/2} \mathbf{F}_{i,j-1/2,k}^\theta}{r_i \sin \theta_j \Delta \theta} \\ & - \frac{\mathbf{F}_{i,j,k+1/2}^\phi - \mathbf{F}_{i,j,k-1/2}^\phi}{r_i \sin \theta_j \Delta \phi} + S_{i,j,k}. \end{aligned} \quad (\text{F.17})$$

3.3 Spherically symmetric motion of gas with velocity $\beta = v/c$ corresponds to four-velocity $u^\alpha = (W, W\beta, 0, 0)$ in spherical coordinates ($c = 1$). Rest-mass conservation gives

$$\frac{1}{r^2} \frac{d}{dt} (r^2 \varrho W) + W \varrho \frac{\partial \beta}{\partial r} = 0, \quad (\text{F.18})$$

where d/dt is the convective derivative defined by $\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \beta \frac{\partial}{\partial r}$. Similarly, $\nabla_\mu T_\alpha^\mu = 0$ yields

$$\frac{1}{r^2} \frac{d}{dt} (r^2 W h u_\alpha) + W h u_\alpha \frac{\partial \beta}{\partial r} + \partial_\alpha p = 0, \quad (\text{F.19})$$

where $h = e + p$. This gives two independent equations ($\alpha = 0, 1$). Choose $\nabla_\mu T_1^\mu = 0$ as one equation, and the projection $u^\alpha \nabla_\mu T_\alpha^\mu = 0$ as the other. Show that conservation of momentum and energy may be expressed by

$$\frac{1}{r^2} \frac{d}{dt} (r^2 W^2 h \beta) = -\frac{\partial p}{\partial r} - W^2 h \beta \frac{\partial \beta}{\partial r} \quad (\text{F.20})$$

$$\frac{1}{r^2} \frac{d}{dt} (r^2 W h) = W \frac{dp}{dt} - W h \frac{\partial \beta}{\partial r}. \quad (\text{F.21})$$

Apply equations (F.18) and (F.20) to the blast (the gas between FS and RS), and make the approximation $\partial \beta / \partial r = 0$, i.e.

$$W(t, r) = \Gamma(t), \quad r_r < r < r_f, \quad (\text{F.22})$$

where $r_r(t)$ and $r_f(t)$ are the instantaneous radii of RS and FS, respectively. Then integration of equations (F.18) and (F.20) over r between RS and FS (at $t = \text{const}$) yields

$$\frac{\Gamma}{r^2} \frac{d}{dr} (r^2 \Sigma \Gamma) = \varrho_r (\beta - \beta_r) \Gamma^2 + \frac{1}{4} \varrho_f \quad (\text{F.23})$$

$$\frac{1}{r^2} \frac{d}{dr} (r^2 H \Gamma^2) = h_r (\beta - \beta_r) \Gamma^2 + p_r \quad (\text{F.24})$$

$$\frac{\Gamma}{r^2} \frac{d}{dr} (r^2 H \Gamma) = \Gamma^2 \frac{dP}{dr} + (h_r - p_r)(\beta - \beta_r) \Gamma^2 + \frac{3}{4} p_f, \quad (\text{F.25})$$

where $\Sigma \equiv \int_{r_r}^{r_f} \varrho dr$, $H \equiv \int_{r_r}^{r_f} h dr$, and $P \equiv \int_{r_r}^{r_f} p dr$. In the derivation of these equations we used the identity for a function $f(t, r)$ and $F(t) = \int_{r_r}^{r_f} f(t, r) dr$,

$$\int_{r_r(t)}^{r_f(t)} \frac{df}{dt} dr = \frac{dF}{dt} - f_f(\beta_f - \beta) - f_r(\beta - \beta_r). \quad (\text{F.26})$$

Here $f_f(t) \equiv f(t, r_f[t])$ and $f_r(t) \equiv f(t, r_r[t])$; $\beta_r = dr_r/dt$ and $\beta_f = dr_f/dt$ are the velocities of RS and FS in the lab frame. The relativistic blast is a very thin shell, $r_f - r_r \sim r/\Gamma^2 \ll r$, and we used $r_f \approx r_r \approx r$ when calculating the integrals. In the integrated equations we took into account that $\Gamma \gg 1$. Then the jump conditions at the FS give $\beta_f - \beta = 1/4\Gamma^2$ and $h_f = 4p_f \gg \varrho_f$. The convective derivative d/dt has been replaced by $\beta d/dr \approx d/dr$ and $\Gamma^2 \beta$ by Γ^2 in the second equation.

Problems in Chapter 4

4.1 The perturbed orbital equation can be written as

$$\begin{aligned} \frac{d^2 u_1}{d\phi^2} + u_1 &= \frac{3G^2 M^2}{L^2} (1 + e \cos \phi)^2 \\ &= \frac{3G^2 M^2}{L^2} [(1 + e^2/2) + 2e \cos \phi + (e^2/2) \cos 2\phi]. \end{aligned} \quad (\text{F.27})$$

This equation can be solved by means of the identity

$$\frac{d^2}{d\phi^2} (\phi \cos \phi) + \phi \sin \phi = 2 \cos \phi. \quad (\text{F.28})$$

A solution to the perturbed equation is then given by

$$u_1 = \frac{3G^2 M^2}{L^2} \left[1 + e^2/2 + e\phi \sin \phi - \frac{e^2}{6} \cos 2\phi \right]. \quad (\text{F.29})$$

The first term is simply a constant offset, and the third term oscillates around zero. The second term represents a secular perturbation which accumulates over the orbits. The full solution can therefore be written as

$$u = 1 + e \cos \phi + \frac{3G^2 M^2 e}{L^2} \phi \sin \phi. \quad (\text{F.30})$$

This expression can be rewritten as an equation for an ellipse with an angular period deviating from 2π

$$u = 1 + e \cos[(1 - \Delta)\phi], \quad (\text{F.31})$$

where we have defined

$$\Delta = \frac{3G^2M^2}{L^2}. \quad (\text{F.32})$$

We have therefore found that a planet suffers a perihelion advance each orbit by an angle

$$\Delta\phi = 2\pi\Delta = \frac{6\pi G^2 M^2}{L^2}. \quad (\text{F.33})$$

4.2 The Christoffel symbols for Schwarzschild are

$$\begin{aligned} \Gamma_{tr}^t &= \frac{GM}{r(r-2GM)} & \Gamma_{tt}^r &= \frac{GM}{r^3}(r-2GM) & \Gamma_{rr}^r &= -\frac{GM}{r(r-2GM)} \\ \Gamma_{r\theta}^\theta &= \frac{1}{r} & \Gamma_{\theta\theta}^r &= -(r-2GM) & \Gamma_{r\phi}^\phi &= \frac{1}{r} \\ \Gamma_{\phi\phi}^r &= -(r-2GM)\sin^2\theta & \Gamma_{\phi\phi}^\theta &= -\sin\theta\cos\theta & \Gamma_{\theta\phi}^\phi &= \frac{\cos\theta}{\sin\theta}. \end{aligned} \quad (\text{F.34})$$

The geodesics equations give then the following relations

$$\frac{d^2t}{d\lambda^2} + \frac{2GM}{r(r-2GM)} \frac{dt}{d\lambda} \frac{dr}{d\lambda} = 0 \quad (\text{F.35})$$

$$\begin{aligned} \frac{d^2r}{d\lambda^2} + \frac{GM}{r^3}(r-2GM) \left(\frac{dt}{d\lambda} \right)^2 - \frac{GM}{r(r-2GM)} \left(\frac{dr}{d\lambda} \right)^2 \\ -(r-2GM) \left[\left(\frac{d\theta}{d\lambda} \right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda} \right)^2 \right] = 0 \end{aligned} \quad (\text{F.36})$$

$$\frac{d^2\theta}{d\lambda^2} + \frac{2}{r} \frac{dr}{d\lambda} \frac{d\theta}{d\lambda} - \sin\theta\cos\theta \left(\frac{d\phi}{d\lambda} \right)^2 = 0 \quad (\text{F.37})$$

$$\frac{d^2\phi}{d\lambda^2} + \frac{2}{r} \frac{dr}{d\lambda} \frac{d\phi}{d\lambda} + 2\frac{\cos\theta}{\sin\theta} \frac{d\theta}{d\lambda} \frac{d\phi}{d\lambda} = 0. \quad (\text{F.38})$$

The third equation shows that a particle with initially $\dot{\theta} = 0$ in the equatorial plane will stay in the equatorial plane.

Problems in Chapter 6

6.1 Particle density, energy density and pressure are given for an ideal Fermi gas

$$n = \frac{8\pi}{(2\pi\hbar)^3} \int_0^{p_F} p^2 dp \quad (\text{F.39})$$

$$\varrho = \frac{8\pi}{(2\pi\hbar)^3} \int_0^{p_F} \sqrt{p^2 + m^2} p^2 dp \quad (\text{F.40})$$

$$P = \frac{1}{3} \frac{8\pi}{(2\pi\hbar)^3} \int_0^{p_F} \frac{p^2}{\sqrt{p^2 + m^2}} p^2 dp. \quad (\text{F.41})$$

Chemical equilibrium between protons, neutrons and electrons requires for the chemical potentials

$$\mu_n = \mu_p + \mu_e , \quad (\text{F.42})$$

where

$$\mu = \epsilon_F = \sqrt{p_F^2 + m^2} = \sqrt{\Lambda^2 n^{2/3} + m^2} \quad (\text{F.43})$$

with the definition $\Lambda = (3\pi^2\hbar)^{1/3}$. From the chemical equilibrium we obtain

$$\frac{n_p}{n_n} = \frac{1}{8} \left[\frac{1 + \frac{2(m_n^2 - m_p^2 - m_e^2)}{\Lambda^2 n_n^{2/3}} + \frac{(m_n^2 - m_p^2)^2 - 2m_e^2(m_n^2 + m_p^2) + 4m_e^4}{\Lambda^4 n_n^{4/3}}}{1 + m_n^2/\Lambda^2 n_n^{4/3}} \right]^{3/2} . \quad (\text{F.44})$$

With the mass difference $Q = m_n - m_p$, and since $Q \ll m_n$ and $m_e \ll m_n$, we can simplify the expression

$$\frac{n_p}{n_n} = \frac{1}{8} \left[\frac{1 + (4Q/m_n)(\varrho_0/m_n n_n)^{2/3} + 4[(Q^2 - m_e^2)/m_n^2](\varrho_0/m_n n_n)^{4/3}}{1 + (\varrho_0/m_n n_n)^{2/3}} \right]^{3/2} \quad (\text{F.45})$$

where the density $\varrho_0 = m_n^4/\Lambda^3 = 6.11 \times 10^{15} \text{ g/cm}^3$ is a characteristic density. For the Fermi momentum of the electrons one obtains

$$\begin{aligned} p_{F,e}^2 &= \Lambda^2 n_e^{2/3} = m_n^2 (m_n n_n / \varrho_0)^{2/3} (n_p/n_n)^{2/3} \\ &= \frac{(m_n^2/4)(m_n n_n / \varrho_0)^{4/3} + Q m_n (m_n n_n / \varrho_0)^{2/3} + Q^2 - m_e^2}{1 + (m_n n_n / \varrho_0)^{2/3}} . \end{aligned} \quad (\text{F.46})$$

Since the maximal momentum of the electron in the classical neutron decay is $p_{\max} = \sqrt{Q^2 - m_e^2} = 1.19 \text{ MeV}$, in neutron stars we have generally $p_{F,e} \gg p_{\max}$.

For small neutron densities n_n , the proton–neutron ratio decreases with increasing number density, until it reaches a minimum at the density

$$\varrho_{\min} \simeq \varrho_0 \left(\frac{4(Q^2 - m_e^2)}{m_n^2} \right)^{3/4} = 1.28 \times 10^{-4} \varrho_0 = 7.8 \times 10^{11} \text{ g cm}^{-3} . \quad (\text{F.47})$$

Beyond this density, n_p/n_n increases and goes asymptotically to the value of 1/8. At nuclear densities, one finds typical values of $n_p/n_n \simeq 0.01$ and $p_{F,e} \simeq 100 \text{ MeV}$. For much higher densities, muons are created, since $p_{F,e} > 105 \text{ MeV}$.

6.2 The maximal mass of a neutron star consisting of noninteracting neutrons is $M_{\max} = 0.71 M_\odot$ with a central density $\varrho_c = 4 \times 10^{15} \text{ g/cm}^3$ and a radius $R = 9.6 \text{ km}$.

6.3 Solvers for the TOV equations can easily be found on the web.

6.6 From the expressions for the Kepler problem one obtains for the time derivative $\dot{\phi}$, which follows from the angular momentum

$$L = \frac{M_1 M_2}{M_1 + M_2} r^2 \dot{\phi} \quad (\text{F.48})$$

and hence together with the expression for L

$$\dot{\phi} = \frac{(M_1 + M_2) a \sqrt{1 - e^2}}{r^2}. \quad (\text{F.49})$$

In addition, the expression for r implies therefore

$$\dot{r} = e \sin \phi \sqrt{\frac{M_1 + M_2}{a(1 - e^2)}}. \quad (\text{F.50})$$

With these relations one can derive the first and second time derivatives of the moments of inertia

$$I_{xx} = M_1 x_1^2 + M_2 x_2^2 = \frac{M_1 M_2}{M_1 + M_2} r^2 \cos^2 \phi \quad (\text{F.51})$$

$$I_{yy} = \frac{M_1 M_2}{M_1 + M_2} r^2 \sin^2 \phi \quad (\text{F.52})$$

$$I_{xy} = \frac{M_1 M_2}{M_1 + M_2} r^2 \sin \phi \cos \phi \quad (\text{F.53})$$

$$I = I_{xx} + I_{yy} = \frac{M_1 M_2}{M_1 + M_2} r^2. \quad (\text{F.54})$$

the following expressions

$$\dot{I}_{xx} = -\frac{2M_1 M_2}{\sqrt{(M_1 + M_2)a(1 - e^2)}} r \cos \phi \sin \phi \quad (\text{F.55})$$

$$\ddot{I}_{xx} = -\frac{2M_1 M_2}{a(1 - e^2)} (\cos 2\phi + e \cos^3 \phi) \quad (\text{F.56})$$

$$\dot{I}_{yy} = \frac{2M_1 M_2}{\sqrt{(M_1 + M_2)a(1 - e^2)}} r (\sin \phi \cos \phi + e \sin \phi) \quad (\text{F.57})$$

$$\ddot{I}_{yy} = \frac{2M_1 M_2}{a(1 - e^2)} (\cos 2\phi + e \cos \phi + e \cos^3 \phi + e^2) \quad (\text{F.58})$$

$$\dot{I}_{xy} = \frac{2M_1 M_2}{\sqrt{(M_1 + M_2)a(1 - e^2)}} r (\cos^2 \phi - \sin^2 \phi + e \cos \phi) \quad (\text{F.59})$$

$$\ddot{I}_{xy} = -\frac{2M_1 M_2}{a(1 - e^2)} (\sin 2\phi + e \sin \phi + e \sin \phi \cos^2 \phi). \quad (\text{F.60})$$

From here we get the third time derivatives

$$\ddot{I}_{xx} = \frac{2M_1 M_2}{a(1-e^2)} (2 \sin 2\phi + 3e \cos^2 \phi \sin \phi) \dot{\phi} \quad (\text{F.61})$$

$$\ddot{I}_{yy} = -\frac{2M_1 M_2}{a(1-e^2)} (2 \sin 2\phi + e \sin \phi + 3e \cos^2 \phi \sin \phi) \dot{\phi} \quad (\text{F.62})$$

$$\ddot{I}_{xy} = -\frac{2M_1 M_2}{a(1-e^2)} (2 \cos 2\phi - e \cos \phi + 3e \cos^3 \phi) \dot{\phi} \quad (\text{F.63})$$

$$\ddot{I} = -\frac{2M_1 M_2}{a(1-e^2)} e \sin \phi \dot{\phi}. \quad (\text{F.64})$$

These quantities determine the energy loss

$$-\frac{dE}{dt} = \frac{1}{5} \left[\ddot{I}_{ik} \ddot{I}_{ik} - \frac{1}{3} (\ddot{I})^2 \right] = \frac{1}{5} \left[(\ddot{I}_{xx})^2 + (\ddot{I}_{yy})^2 + 2(\ddot{I}_{xy})^2 - \frac{1}{3} (\ddot{I})^2 \right]. \quad (\text{F.65})$$

6.9 The Crab Nebula is a unique cosmic lab with an extremely broad spectrum of nonthermal radiation (radio and optical emission is strongly polarized). The spectrum extends through 20 decades in frequency space, ranging from radio wavelengths to gamma-ray emission. It is commonly assumed that the synchrotron nebula is powered by electrons and positrons generated by the central pulsar and terminated by a standing reverse shock at a distance of about 0.1 pc from the pulsar (Fig. 1.8). A discussion of the global spectrum of the Crab Nebula (Fig. F.1) can be found in Aharonian and Atoyan [24] and Aharonian et al. [25].

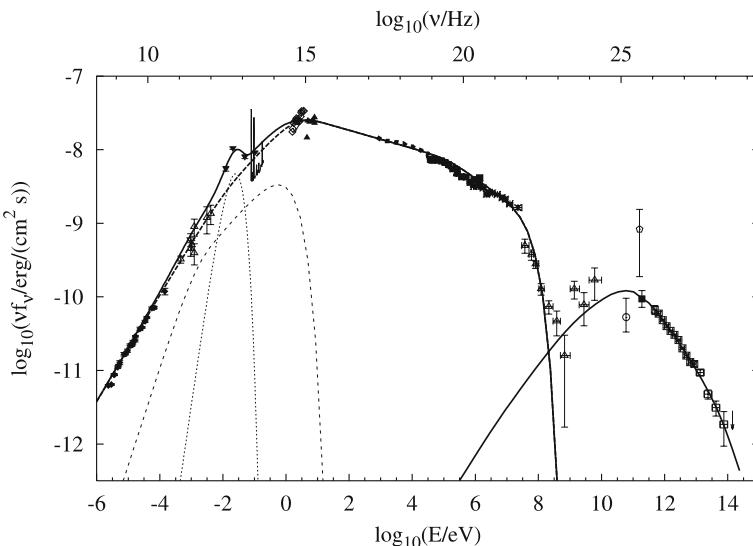


Fig. F.1. Energy spectrum of the Crab Nebula compiled from the literature. The *solid* and *dashed curves* correspond to synchrotron and inverse Compton emission, respectively. Figure adapted from [25]

6.10 We start with the Boltzmann equation for massless particles (photons or neutrinos)

$$p^\alpha \frac{\partial f}{\partial x^\alpha} + \Gamma_{\beta\gamma}^\alpha p^\beta p^\gamma \frac{\partial f}{\partial p^\alpha} = \left(\frac{df}{d\tau} \right)_{\text{coll}}, \quad (\text{F.66})$$

where f is the invariant neutrino (photon) distribution function, p^α the neutrino four-momentum and Γ the Christoffel symbols for the metric of a neutron star. Since the collision term is only known in a comoving frame, we transform to comoving momenta p^α

$$p^a e_a^\alpha \frac{\partial f}{\partial x^\alpha} + e_a^\alpha \omega_{bc}^a p^b p^c \frac{\partial f}{\partial p^\alpha} = \left(\frac{df}{d\tau} \right)_{\text{coll}}, \quad (\text{F.67})$$

with basis vectors \mathbf{e}_a for the comoving observer. The connection coefficients are those derived in Sect. 4.2

$$\omega_{00}^1 = \exp(-\lambda) (\partial_r \Phi) \quad (\text{F.68})$$

$$\omega_{22}^1 = -\frac{\exp(-\lambda)}{r} = \omega_{33}^1 = -\omega_{21}^2 = -\omega_{31}^3 \quad (\text{F.69})$$

$$\omega_{33}^2 = -\frac{\cot\theta}{r} = -\omega_{32}^3. \quad (\text{F.70})$$

The neutrino four-momentum can be parametrized by

$$p^a = \hat{E} \left(1, \mu, \sqrt{1-\mu^2} \cos \chi, \sqrt{1-\mu^2} \sin \chi \right), \quad (\text{F.71})$$

where μ is the cosine of the angle between the neutrino momentum and the radial direction, \hat{E} is the neutrino energy in the comoving frame. With these quantities, the Boltzmann equation is simply given by

$$\begin{aligned} \hat{E} e_0^t \frac{\partial f}{\partial t} + \hat{E} e_1^r \frac{\partial f}{\partial r} - \hat{E} \mu e_0^t \frac{\partial f}{\partial r} \hat{E}^2 \mu \omega_{00}^1 \frac{\partial f}{\partial \hat{E}} \\ - \hat{E} (1 - \mu^2) (\omega_{00}^1 + \omega_{22}^1) \frac{\partial f}{\partial \mu} = \left(\frac{df}{d\tau} \right)_{\text{coll}}. \end{aligned} \quad (\text{F.72})$$

As with photon distributions, one defines the n th moments

$$M_n = \frac{1}{2} \int_{-1}^1 d\mu \mu^n f, \quad Q_n = \frac{1}{2} \int_{-1}^1 d\mu \mu^n \left(\frac{df}{d\tau} \right)_{\text{coll}}. \quad (\text{F.73})$$

They satisfy the following evolution equations

$$\begin{aligned} \hat{E} \left(e_0^t \frac{\partial M_0}{\partial t} + e_1^r \frac{\partial M_1}{\partial r} \right) - \hat{E}^2 \omega_{00}^2 \frac{\partial M_1}{\partial \hat{E}} \\ - 2\hat{E} (\omega_{00}^1 + \omega_{22}^1) M_1 = Q_0 \end{aligned} \quad (\text{F.74})$$

$$\begin{aligned} \hat{E} \left(e_0^t \frac{\partial M_1}{\partial t} + e_1^r \frac{\partial M_2}{\partial r} \right) - \hat{E}^2 \omega_{00}^2 \frac{\partial M_2}{\partial \hat{E}} \\ + \hat{E} (\omega_{00}^1 + \omega_{22}^1) (M_0 - 3M_2) = Q_1. \end{aligned} \quad (\text{F.75})$$

We now introduce the number density N_ν , the number flux F_ν and the number source term S_N , together with the mean neutrino energy density J_ν , energy flux H_ν , pressure P_ν and energy source term S_E by means of the definitions

$$N_\nu = \frac{1}{2\pi^2} \int_0^\infty M_0 \hat{E}^2 d\hat{E} \quad (\text{F.76})$$

$$F_\nu = \frac{1}{2\pi^2} \int_0^\infty M_1 \hat{E}^2 d\hat{E} \quad (\text{F.77})$$

$$S_N = \frac{1}{2\pi^2} \int_0^\infty Q_0 \hat{E} d\hat{E} \quad (\text{F.78})$$

$$J_\nu = \frac{1}{2\pi^2} \int_0^\infty M_0 \hat{E}^3 d\hat{E} \quad (\text{F.79})$$

$$H_\nu = \frac{1}{2\pi^2} \int_0^\infty M_1 \hat{E}^3 d\hat{E} \quad (\text{F.80})$$

$$P_\nu = \frac{1}{2\pi^2} \int_0^\infty M_2 \hat{E}^3 d\hat{E} \quad (\text{F.81})$$

$$S_E = \frac{1}{2\pi^2} \int_0^\infty Q_0 \hat{E}^2 d\hat{E}. \quad (\text{F.82})$$

After integration over the neutrino energy and by using the continuity equation, one recovers the neutrino (photon) transport equations

$$\frac{\partial[N_\nu/n_B]}{\partial t} + \frac{\partial[4\pi r^2 \exp \Phi F_\nu]}{\partial r} = \exp \Phi \frac{S_N}{n_B} \quad (\text{F.83})$$

$$\begin{aligned} \frac{\partial[J_\nu/n_B]}{\partial t} &+ P_\nu \frac{\partial[1/n_B]}{\partial t} \\ &+ \exp(-\Phi) \frac{\partial[4\pi r^2 \exp(2\Phi) H_\nu]}{\partial r} = \exp \Phi \frac{S_E}{n_B}. \end{aligned} \quad (\text{F.84})$$

The flux of energy $L(r)$ per unit time through a spherical shell at distance r from the center is proportional to the gradient of the temperature

$$L(r) = -4\pi r^2 \kappa(r) \frac{\partial[\exp(\Phi)T]}{\partial r} \exp(-\Phi) \sqrt{1 - \frac{2m(r)}{r}}, \quad (\text{F.85})$$

where the factor $\exp(-\Phi)\sqrt{1 - 2m(r)/r}$ corresponds to the relativistic correction of the time-scale (redshift) and the shell thickness. For this purpose, it is useful to introduce the shell volume $A(r)$ defined by

$$\frac{\partial A}{\partial r} = \frac{4\pi r^2 n}{\sqrt{1 - 2m(r)/r}}, \quad (\text{F.86})$$

where $n(r)$ is the baryon number density. With this shell variable, one can write the energy balance and the thermal energy transport as

$$\frac{\partial}{\partial A} [L(A) \exp(2\Phi)] = -\frac{1}{n} \left[\epsilon_v \exp(2\Phi) + c_V \frac{\partial[T \exp \Phi]}{\partial t} \right] \quad (\text{F.87})$$

$$\frac{\partial}{\partial A} [T(A) \exp(\Phi)] = -\frac{3\kappa\varrho}{4acT^3} \frac{L \exp \Phi}{4\pi r^2}. \quad (\text{F.88})$$

These equations are supplemented by the stellar structure equations for the mass distribution $m(A)$ and the potential $\Phi(A)$

$$\frac{\partial m}{\partial A} = \frac{\varrho}{n} \sqrt{1 - \frac{2m(A)}{r}} \quad (\text{F.89})$$

$$\frac{\partial \Phi}{\partial A} = \frac{4\pi r^3 P + m(A)}{4\pi r^2 n} \frac{1}{\sqrt{1 - 2m(A)/r}}. \quad (\text{F.90})$$

The pressure profile follows from the hydrodynamical equilibrium (see Sect. 2.7)

$$\frac{\partial P}{\partial A} = -(\varrho + P) \frac{\partial \Phi}{\partial A}. \quad (\text{F.91})$$

6.11 For systems consisting of a radio pulsar and a white dwarf, it is extremely difficult to measure the periastron advance, since these systems are highly circular. Then also time dilation and gravitational redshift are difficult to measure. The only relativistic effect, which can be used, is the Shapiro time delay. If the range and shape of the Shapiro delay can be measured, this gives the mass M_c of the companion (WD) and the inclination $\sin i$. Together with the mass function, this provides an accurate measurement of the neutron star mass in such systems. With this method, the two post-Keplerian parameters have been measured for the millisecond pulsar PSR J1909–3744 ($P = 2.95$ ms, $P_b = 1.533449$ d, projected semimajor axis $a \sin i = 1.89799$ lt-s) [216]: (i) the range parameter $r = GM_c/c^3 = (1.004 \pm 0.011)$ μ s and (ii) the shape parameter $s = \sin i = 0.99822 \pm 0.00011$. This gives the mass of the white dwarf $M_c = (0.2038 \pm 0.0022) M_\odot$ and the orbital inclination $i = (86.58 \pm 0.11)$ degrees. Therefore, we get the neutron star mass $M_n = (1.438 \pm 0.024) M_\odot$.

The eccentricity of the system is extremely low, $e = 1.35 \times 10^{-7}$. This corresponds to $\dot{\omega}$ predicted by GR in the range of 0.14 deg yr $^{-1}$. Since this mass of the neutron star in a recycled system is only slightly higher than the masses observed for other systems, it appears that the production of a millisecond pulsar is possible with the accretion of less than $0.2 M_\odot$. Most of the mass would be lost from the system.

Problems in Chapter 7

7.1 Since $\sqrt{\gamma} = \exp(\psi + \mu_2 + \mu_3)$, we find

$$\begin{aligned} \alpha \operatorname{Div} \left[\frac{\nabla \psi}{\alpha} \right] &= \frac{1}{\sqrt{\gamma}} \left\{ \partial_A \left[\frac{\sqrt{\gamma}}{\alpha} g^{AB} \partial_B \psi \right] \right\} \\ &= \frac{\alpha}{\sqrt{\gamma}} \left\{ \frac{\exp \mu_3}{\alpha} \partial_2 [\exp(\psi - \mu_2) \partial_2 \psi] + \frac{\exp \mu_2}{\alpha} \partial_3 [\exp(\psi - \mu_3) \partial_3 \psi] \right. \\ &\quad + \frac{\exp(\psi - \mu_2)}{\alpha} (\partial_2 \psi) \exp \mu_3 (\partial_2 \mu_3) \\ &\quad + \frac{\exp(\psi - \mu_3)}{\alpha} (\partial_3 \psi) \exp \mu_2 (\partial_3 \mu_2) \\ &\quad - \frac{\exp(\psi - \mu_2 + \mu_3)}{\alpha^2} (\partial_2 \psi)(\partial_2 \alpha) \\ &\quad \left. - \frac{\exp(\psi - \mu_3 + \mu_2)}{\alpha^2} (\partial_3 \psi)(\partial_3 \alpha) \right\} \\ &= \frac{1}{R} \nabla_A [R \Psi_A] + \Psi_3 \nabla_3 \mu_2 + \Psi_2 \nabla_2 \mu_3 - \frac{1}{\alpha} (\nabla_A \Psi) (\nabla_A \alpha). \end{aligned} \quad (\text{F.92})$$

7.2 The components of the Riemann tensor in the orthonormal frame follow from the curvature two-form

$$\Omega_a^b = \frac{1}{2} R_{bcd}^a \Theta^c \wedge \Theta^d, \quad (\text{F.93})$$

with the following expressions

$$R_{1010} = -\Psi_2 (\nabla_2 v) - \Psi_3 (\nabla_3 v) - \frac{1}{4} \frac{R^2}{\alpha^2} (\nabla \omega \cdot \nabla \omega) \quad (\text{F.94})$$

$$R_{2020} = -\frac{1}{\alpha} \nabla_2 (\nabla_2 \alpha) - (\nabla_3 v) (\nabla_3 \mu_2) + \frac{3}{4} \frac{R^2}{\alpha^2} (\nabla_2 \omega)^2 \quad (\text{F.95})$$

$$R_{3030} = -\frac{1}{\alpha} \nabla_3 (\nabla_3 \alpha) - (\nabla_2 v) (\nabla_2 \mu_3) + \frac{3}{4} \frac{R^2}{\alpha^2} (\nabla_3 \omega)^2 \quad (\text{F.96})$$

$$R_{3020} = -\frac{1}{\alpha} \nabla_3 (\nabla_2 \alpha) + (\nabla_3 v) (\nabla_2 \mu_3) + \frac{3}{4} \frac{R^2}{\alpha^2} (\nabla_3 \omega) (\nabla_2 \omega) \quad (\text{F.97})$$

$$R_{1212} = \frac{1}{R} \nabla_2 (R \Psi_2) + \Psi_3 (\nabla_3 \mu_2) + \frac{1}{4} \frac{R^2}{\alpha^2} (\nabla_2 \omega)^2 \quad (\text{F.98})$$

$$R_{1313} = \frac{1}{R} \nabla_3 (R \Psi_3) + \Psi_2 (\nabla_2 \mu_3) + \frac{1}{4} \frac{R^2}{\alpha^2} (\nabla_3 \omega)^2 \quad (\text{F.99})$$

$$\begin{aligned} R_{2323} &= \exp(-\mu_2) \nabla_3 [\exp(\mu_2) \nabla_3 \mu_2] \\ &\quad + \exp(-\mu_3) \nabla_2 [\exp(\mu_3) \nabla_2 \mu_3] = \Delta(\mu_2, \mu_3) \end{aligned} \quad (\text{F.100})$$

$$R_{1213} = \frac{1}{R} \nabla_3 (R \Psi_2) - \Psi_3 (\nabla_2 \mu_3) + \frac{1}{4} \frac{R^2}{\alpha^2} (\nabla_2 \omega) (\nabla_3 \omega) \quad (\text{F.101})$$

$$\begin{aligned} R_{1023} = & -\frac{1}{2} \exp(-\mu_3)/\alpha \nabla_2[(R \exp(\mu_3) \nabla_3 \omega] \\ & + \frac{1}{2} \exp(-\mu_2) \nabla_3[(R \exp(\mu_2)/\alpha) \nabla_2 \omega] \end{aligned} \quad (\text{F.102})$$

$$\begin{aligned} R_{2012} = & \frac{R}{\alpha} (\nabla_2 \omega)(\Psi_2 - \nabla_2 v/2) + \frac{1}{2\alpha} \nabla_2[R \nabla_2 \omega] \\ & + \frac{R}{2\alpha} (\nabla_3 \omega)(\nabla_3 \mu_2) \end{aligned} \quad (\text{F.103})$$

$$R_{2013} = \frac{1}{2\alpha} \nabla_2[R \nabla_3 \omega] + \frac{R}{2\alpha} (\nabla_2 \omega)[2\Psi_3 - \nabla_3 \mu_2 - \nabla_3 v] \quad (\text{F.104})$$

$$\begin{aligned} R_{3013} = & \frac{1}{2\alpha} \nabla_3[R \nabla_3 \omega] + \frac{R}{\alpha} (\nabla_3 \omega)[\Psi_3 - \nabla_3 v/2] \\ & + \frac{R}{2\alpha} (\nabla_2 \omega)(\nabla_2 \mu_3). \end{aligned} \quad (\text{F.105})$$

The following components vanish identically

$$R_{1012} = 0 = R_{1013} = R_{1002} = R_{1003} = R_{2023} = R_{3023} = R_{1223} = R_{1323}. \quad (\text{F.106})$$

From these curvature, you can directly obtain the Ricci tensors.

7.4 For the calculation of the hydrodynamical equilibrium in the metric field of a rotating star

$$P_{,A} + \frac{P}{\sqrt{-g}} \partial_A \left[\sqrt{-g} \right] = T^{BC} \partial_B(g_{CA}) - T^{\mu\nu} \Gamma_{\mu\nu}^B g_{BA}, \quad (\text{F.107})$$

we need the Christoffel symbols ($A = 2, 3$)

$$\Gamma_{00}^A = -\frac{1}{2} \exp(-2\mu) \partial_A [\omega^2 \exp 2\Psi - \alpha^2] \quad (\text{F.108})$$

$$\Gamma_{11}^A = -(\partial_A \Psi) \exp(2\Psi - 2\mu), \quad \Gamma_{22}^A = \partial_A \mu \quad (\text{F.109})$$

$$\Gamma_{01}^A = \frac{1}{2} \exp(-2\mu) \partial_A (\omega \exp \Psi), \quad \Gamma_{33}^A = -\partial_A \mu. \quad (\text{F.110})$$

By inserting these relations into the above equation, we obtain

$$\begin{aligned} P_{,A} = & -P \left[\partial_A v + \partial_A \Psi + \frac{1}{2} \exp(-2v) \partial_A (\omega^2 \exp 2\Psi - \exp 2v) \right. \\ & \left. - \omega \partial_A (\omega \exp 2\Psi) - \exp 2\Psi (\exp -2\Psi - \omega^2 \exp -2v) \right] \\ & - (\varrho + P)(u')^2 \left[-\frac{1}{2} \exp(-2v) \partial_A (\omega^2 \exp 2\Psi - \exp 2v) \right. \\ & \left. + \Omega \partial_A (\omega \exp 2\Psi) - \Omega^2 \exp 2\Psi \partial_A \Psi \right] \end{aligned}$$

$$\begin{aligned}
&= -(\varrho + P)\partial_A \ln \alpha \\
&\quad -(\varrho + P)(u^t)^2 \exp 2\Psi \left[(\omega - \Omega)^2 \partial_A v - (\omega - \Omega) \partial_A \omega \right. \\
&\quad \left. - (\omega - \Omega)^2 \partial_A \Psi \right]. \tag{F.111}
\end{aligned}$$

In the last bracket we add and subtract the term $(\omega - \Omega) \partial_A \Omega$

$$\begin{aligned}
P_{,A} &= -(\varrho + P)\partial_A \ln \alpha \\
&\quad -(\varrho + P) \frac{\gamma^2}{\alpha^2} \exp 2\Psi \left[(\omega - \Omega)^2 \partial_A v - (\omega - \Omega) \partial_A (\omega - \Omega) \right. \\
&\quad \left. - (\omega - \Omega)^2 \partial_A \Psi - (\omega - \Omega) \partial_A \Omega \right] \\
&= -(\varrho + P)\partial_A \ln \alpha \\
&\quad + \frac{1}{2}(\varrho + P) \frac{\gamma^2}{\alpha^2} \exp 2v \partial_A [\exp(2\Psi - 2v)(\omega - \Omega)^2] \\
&\quad + (\varrho + P) \frac{\gamma^2}{\alpha^2} \exp 2\Psi (\omega - \Omega) \partial_A \Omega. \tag{F.112}
\end{aligned}$$

Since the three-velocity is given by

$$V^2 = \exp(2\Psi - 2v)(\omega - \Omega)^2 \tag{F.113}$$

with the Lorentz factor $\gamma = 1/\sqrt{1 - V^2}$ and $u^t = \gamma/\alpha$, we obtain

$$\begin{aligned}
P_{,A} &= -(\varrho + P)\partial_A \ln \alpha \\
&\quad + \frac{1}{2}(\varrho + P)\gamma^2 \partial_A [V^2 - 1] \\
&\quad + (\varrho + P)(u^t)^2 \exp 2\Psi (\omega - \Omega) \partial_A \Omega. \tag{F.114}
\end{aligned}$$

Now, we use the identity

$$(u^t)^2 \exp 2\Psi (\omega - \Omega) = -u^t [g_{01}u^t + g_{11}u^1] = -u^t u_\phi. \tag{F.115}$$

With this, the equilibrium condition can be written as

$$P_{,A} = (\varrho + P) \left[\partial_A \ln \left(\frac{\gamma}{\alpha} \right) - u^t u_\phi \partial_A \Omega \right]. \tag{F.116}$$

Problems in Chapter 8

8.1 From the discussion in Sect. 4.4, we get the two expressions for bound orbits

$$\tau = \frac{1}{L} \int \frac{d\phi}{u^2} = \frac{1}{L} \int \frac{d\phi}{d\chi} \frac{d\chi}{u^2} \tag{F.117}$$

and

$$t = \frac{E}{L} \int \frac{d\phi}{d\chi} \frac{d\chi}{u^2(1 - 2Mu)} . \quad (\text{F.118})$$

Using the explicit solution, we can bring this to the form

$$\tau = \frac{l^{3/2}}{\sqrt{M}} \sqrt{1 - \mu(3 + e^2)} \int_{\chi}^{\pi} \frac{d\chi}{(1 + e \cos \chi)^2 \sqrt{1 - 2\mu(3 + e \cos \chi)}} \quad (\text{F.119})$$

and

$$t = \frac{l^{3/2}}{\sqrt{M}} \sqrt{(2\mu - 1)^2 - 4\mu^2 e^2} \\ \times \int_{\chi}^{\pi} \frac{d\chi}{(1 + e \cos \chi)^2 \sqrt{1 - 2\mu(3 + e \cos \chi)} \sqrt{1 - 2\mu(1 + e \cos \chi)}} . \quad (\text{F.120})$$

With the definition of the Newtonian period

$$P = \sqrt{\frac{4\pi^2 l^3}{(1 - e^2)^3 GM}} , \quad (\text{F.121})$$

the factors in front of the integrals can be written as

$$\frac{1}{2\pi} P (1 - e^2)^{3/2} \sqrt{1 - \mu(3 + e^2)} \quad (\text{F.122})$$

and

$$\frac{1}{2\pi} P (1 - e^2)^{3/2} \sqrt{(2\mu - 1)^2 - 4\mu^2 e^2} . \quad (\text{F.123})$$

In the case $e = 0$, the orbit is a circle with radius r_c given by

$$r_c = l , \quad \mu = M/r_c . \quad (\text{F.124})$$

Angular momentum L and energy E of the orbit are related to the parameter l and e given by

$$L^2 = \frac{Mr_c}{1 - 3M/r_c} , \quad \frac{E^2}{L^2} = \frac{(1 - 2M/r_c)^2}{Mr_c} . \quad (\text{F.125})$$

The first equation gives a quadratic equation for the radius

$$r_c^2 - L^2 r_c/M + 3L^2 = 0 \quad (\text{F.126})$$

with the solutions

$$r_c = \frac{L^2}{2M} \left[1 \pm \sqrt{1 - 12M^2/L^2} \right] . \quad (\text{F.127})$$

Therefore, no circular orbit is possible for $L/M < 2\sqrt{3}$, and for the minimum allowed value of L/M we find

$$r_c = 6M, \quad E^2 = 8/9, \quad L/M = 2\sqrt{3}. \quad (\text{F.128})$$

The larger of the roots locates the minimum of the effective potential curve, while the smaller root locates the maximum in the effective potential. Therefore, the circular orbit with the larger radius will be stable, the orbit of the smaller radius unstable.

The periods for one complete revolution of these circular orbits, measured in proper time and coordinate time t , are

$$\tau_{\text{period}} = P \sqrt{\frac{1 - 3\mu}{1 - 6\mu}} \quad (\text{F.129})$$

and

$$t_{\text{period}} = \frac{P}{\sqrt{1 - 6\mu}}. \quad (\text{F.130})$$

Note that $t_{\text{period}} \rightarrow \infty$ for $r_c \rightarrow 6M$.

8.2 In cosmology, dark energy is a hypothetical form of energy which permeates all of space and has strong negative pressure. According to the theory of relativity, the effect of such a negative pressure is qualitatively similar to a force acting in opposition to gravity at large scales. Invoking such an effect is currently the most popular method for explaining the observations of an accelerating Universe as well as accounting for a significant portion of the missing mass in the Universe.

Two proposed forms for dark energy are the cosmological constant Λ , a constant energy density filling space homogeneously, and quintessence, a dynamic field whose energy density can vary in time and space. Distinguishing between the alternatives requires high-precision measurements of the expansion of the Universe to understand how the speed of the expansion changes over time. The rate of expansion is parameterized by the cosmological equation of state. Measuring the equation of state of dark energy is one of the biggest efforts in observational cosmology today. Other ideas for dark energy have come from string theory, brane cosmology and the holographic principle.

8.3 The Riemann tensors for spherically symmetric spacetimes can be found in Sect. 4.2.1.

8.4 For this purpose, we define two new potentials

$$X = \chi + \omega = \frac{\sqrt{\Delta} + a\delta}{[r^2 + a^2 + a\sqrt{\Delta\delta}]\sqrt{\delta}} \quad (\text{F.131})$$

$$Y = \chi - \omega = \frac{\sqrt{\Delta} - a\delta}{[r^2 + a^2 - a\sqrt{\Delta\delta}]\sqrt{\delta}}. \quad (\text{F.132})$$

The equations

$$-\frac{\mu}{\delta}(\mu_2 + \mu_3)_{,2} + \frac{r - M}{\Delta}(\mu_2 + \mu_3)_{,3} = \frac{2}{(X + Y)^2}(X_{,2}Y_{,3} + Y_{,2}X_{,3}) \quad (\text{F.133})$$

and

$$\begin{aligned} & 2(r - M) \frac{\partial}{\partial r}(\mu_2 + \mu_3) + 2\mu \frac{\partial}{\partial \mu}(\mu_2 + \mu_3) \\ &= \frac{4}{(X + Y)^2}(\Delta X_{,2}Y_{,2} - \delta X_{,3}Y_{,3}) - 3 \frac{M^2 - a^2}{\Delta} - \frac{1}{\delta}. \end{aligned} \quad (\text{F.134})$$

Making use of the solutions for X and Y with their derivatives, one obtains after some calculations the two equations

$$-\frac{\mu}{\delta}(\mu_2 + \mu_3)_{,2} + \frac{r - M}{\Delta}(\mu_2 + \mu_3)_{,3} = \frac{\mu}{\varrho^2 \Delta \delta} [(r - M)(\varrho^2 + 2a^2 \delta) - 2r\Delta] \quad (\text{F.135})$$

and

$$(r - M) \frac{\partial}{\partial r}(\mu_2 + \mu_3) + \mu \frac{\partial}{\partial \mu}(\mu_2 + \mu_3) = 2 - \frac{(r - M)^2}{\Delta} - 2 \frac{rM}{\varrho^2}. \quad (\text{F.136})$$

These two equations are solved by means of the ansatz

$$\boxed{\exp(\mu_2 + \mu_3) = \frac{\varrho^2}{\sqrt{\Delta}}.} \quad (\text{F.137})$$

The solutions for the two meridional metric functions are therefore

$$\boxed{\exp(2\mu_2) = \varrho^2/\Delta, \quad \exp(2\mu_3) = \varrho^2.} \quad (\text{F.138})$$

8.7 The Riemann tensors for the Kerr solution obey the following symmetries

$$R_{1213} = R_{0302}, \quad R_{1330} = R_{1202}, \quad R_{0202} = -R_{1313} \quad (\text{F.139})$$

$$R_{0303} = -R_{1212}, \quad R_{2323} = -R_{0101} = R_{0202} + R_{0303}. \quad (\text{F.140})$$

A straightforward calculation gives the following tetrad components

$$R_{0101} = -R_{2323} = -\frac{Mr}{\varrho^6}(r^2 - 3a^2 \cos^2 \theta) \quad (\text{F.141})$$

$$\begin{aligned} R_{0202} &= -R_{1313} \\ &= \frac{Mr}{\Sigma^2 \varrho^6}(r^2 - 3a^2 \cos^2 \theta)[2(r^2 + a^2)^2 + a^2 \Delta \sin^2 \theta] \end{aligned} \quad (\text{F.142})$$

$$\begin{aligned} R_{0303} &= -R_{1212} \\ &= -\frac{Mr}{\Sigma^2 \varrho^6}(r^2 - 3a^2 \cos^2 \theta)[(r^2 + a^2)^2 + 2a^2 \Delta \sin^2 \theta] \end{aligned} \quad (\text{F.143})$$

$$R_{0123} = \frac{aM \cos \theta}{\varrho^6} (3r^2 - a^2 \cos^2 \theta) \quad (\text{F.144})$$

$$R_{0213} = \frac{aM \cos \theta}{\Sigma^2 \varrho^6} (3r^2 - a^2 \cos^2 \theta)[2(r^2 + a^2)^2 + a^2 \Delta \sin^2 \theta] \quad (\text{F.145})$$

$$R_{0312} = \frac{aM \cos \theta}{\Sigma^2 \varrho^6} (3r^2 - a^2 \cos^2 \theta)[(r^2 + a^2)^2 + 2a^2 \Delta \sin^2 \theta] \quad (\text{F.146})$$

$$\begin{aligned} R_{0302} &= R_{1213} \\ &= -\frac{aM \cos \theta}{\Sigma^2 \varrho^6} (3r^2 - a^2 \cos^2 \theta) 3a\sqrt{\Delta}(r^2 + a^2) \sin \theta \end{aligned} \quad (\text{F.147})$$

$$\begin{aligned} R_{0212} &= R_{1330} \\ &= -\frac{Mr \cos \theta}{\Sigma^2 \varrho^6} (r^2 - 3a^2 \cos^2 \theta) 3a\sqrt{\Delta}(r^2 + a^2) \sin \theta. \end{aligned} \quad (\text{F.148})$$

These components become singular only for $\theta = \pi/2$ and $r = 0$. This is the ring singularity for the Kerr solution.

8.8 Ray-tracers are suitably developed by applying object-oriented methods (C++ classes).

8.9 The equations of motion in the gravitational field of a gravastar can be solved by

$$\frac{dt}{ds} = E \left(\frac{1}{f} - \frac{f\tilde{\omega}^2}{\varrho^2} \right) + L \frac{f\tilde{\omega}}{\varrho^2} \quad (\text{F.149})$$

$$\frac{d\phi}{ds} = -E \frac{f\tilde{\omega}}{\varrho^2} + L \frac{f}{\varrho^2} \quad (\text{F.150})$$

For motion in the equatorial plane we have

$$\left(\frac{dr}{ds} \right)^2 \frac{\exp(2\gamma)}{f} = \frac{E^2}{f} - \frac{f}{\varrho^2} (L - \tilde{\omega}E)^2 \equiv V(\varrho) \quad (\text{F.151})$$

which defines an effective potential $V(\varrho)$. Circular orbits follow from the condition $V = 0 = dV/d\varrho$ which determine the energy and angular momentum

$$E = \frac{\sqrt{f}}{\sqrt{1 - f^2 x^2 / \varrho^2}} \quad (\text{F.152})$$

$$L = E(p + \tilde{\omega}) \quad (\text{F.153})$$

with the definitions

$$p = \varrho^2 \left(-l + \sqrt{l^2 + m - m^2 \varrho^2} \right) / n \quad (\text{F.154})$$

$$l = f\tilde{\omega} \quad (\text{F.155})$$

$$m = \dot{f}/f \quad (\text{F.156})$$

$$n = f - \varrho^2 \dot{f} \quad (\text{F.157})$$

$$\dot{f} \equiv df/d\varrho^2 \quad (\text{F.158})$$

The innermost stable circular orbit (ISCO) follows from $d^2 V/dq^2 = 0$.

Problems in Chapter 9

9.1 The spin evolution of black holes is discussed in [299].

9.2 The evolution of a massive black hole pair is first given by the action of dynamical friction by a uniform background of light stars with isotropic velocity distribution. These black holes form a bounded pair (binary), and the binding energy of the binary increases over time through dynamical friction. If the dynamical friction remains effective until the separation of the two black holes in the binary becomes small enough for the gravitational radiation to shrink the separation further, the two black holes will merge. Such mergings of two massive black holes would emit strong gravitational wave which will be observable through space-based gravitational wave detectors such as LISA. For more details, see [290]

9.3 If the stellar population of the bulge contains black holes formed in the final core collapse of ordinary stars with $M \geq 30M_{\odot}$, then about 25,000 stellar mass black holes should have migrated by dynamical friction into the central parsec of the Milky Way, forming a black hole cluster around the central supermassive black hole. These black holes can be captured by the central black hole when they randomly reach a highly eccentric orbit due to relaxation, either by direct capture (when their Newtonian peribothron is less than four Schwarzschild radii), or after losing orbital energy through gravitational waves. The overall depletion timescale is ~ 30 Gyr, so most of the 25,000 black holes remain in the central cluster today. The presence of this black hole cluster would have several observable consequences. First, the low-mass, old stellar population should have been expelled from the region occupied by the black hole cluster due to relaxation, implying a core in the profile of solar-mass red giants with a radius of ~ 2 pc (i.e. $1'$). The observed central density cusp (which has a core radius of only a few arcseconds) should be composed primarily of young (≤ 1 Gyr) stars. Second, flares from stars being captured by supermassive black holes in other galaxies should be rarer than usually expected because the older stars will have been expelled from the central regions by the black hole clusters of those galaxies. Third, the young (≤ 2 Gyr) stars found at distances $\sim 3\text{--}10$ pc from the Galactic center should be preferentially on highly eccentric orbits. Fourth, if future high-resolution K-band images reveal sources microlensed by the Milky Way's central black hole, then the cluster black holes could give rise to secondary ("planet-like") perturbations on the main event. For more details, see [298].

During five years of Chandra observations, Muno et al. [306] have identified seven X-ray transients located within 23 pc of Sgr A*. These sources each vary in luminosity by more than a factor of 10 and have peak X-ray luminosities greater than 5×10^{33} ergs s $^{-1}$, which strongly suggests that they are accreting black holes or neutron stars. The peak luminosities of the transients are intermediate between those typically considered outburst and quiescence for X-ray binaries. Remarkably, four of these transients lie within only 1 pc of Sgr A*. This implies that, compared to the

numbers of similar systems located between 1 and 23 pc, transients are overabundant by a factor of > 20 per unit stellar mass within 1 pc of Sgr A*. It is likely that the excess transient X-ray sources are low-mass X-ray binaries that were produced, as in the cores of globular clusters, by three-body interactions between binary star systems and either black holes or neutron stars that have been concentrated in the central parsec through dynamical friction. Alternatively, they could be high-mass X-ray binaries that formed among the young stars that are present in the central parsec.

9.6 The light cylinder surfaces are given by the solution of the equation

$$(\Omega_F - \omega)^2 \varpi_L^2 = c^2 \alpha^2 \quad (\text{F.159})$$

for given $\Omega_F = b\Omega_H$ with $b \leq 1$. With the expressions for $\alpha(r, \theta)$, $\omega(r, \theta)$ and $\tilde{\omega}(r, \theta)$ this can be transformed to an implicit equation for $r_L = r_L(\theta)$.

9.7 The general expression for the redshifted energy flux \mathbf{S}_E and the angular momentum flux about the axis of rotation \mathbf{S}_L are given by the energy-momentum tensor

$$\begin{aligned} \mathbf{S}_E &= \frac{1}{4\pi} \left[\alpha(\mathbf{E} \times \mathbf{B}) - \omega(\mathbf{E} \cdot \mathbf{m})\mathbf{E} - \omega(\mathbf{B} \cdot \mathbf{m})\mathbf{B} + \frac{1}{2}\omega(\mathbf{E}^2 + \mathbf{B}^2)\mathbf{m} \right], \\ \mathbf{S}_L &= \frac{1}{4\pi} \left[-(\mathbf{E} \cdot \mathbf{m})\mathbf{E} - (\mathbf{B} \cdot \mathbf{m})\mathbf{B} + \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2)\mathbf{m} \right]. \end{aligned} \quad (\text{F.160})$$

Since the toroidal component of the fluxes are irrelevant, we only need to consider the poloidal components

$$\begin{aligned} \mathbf{S}_L^p &= -\frac{\varpi}{4\pi} |\mathbf{B}_T| \mathbf{B}_p = \frac{I}{2\pi\alpha} \mathbf{B}_p, \\ \mathbf{S}_E^p &= \frac{\alpha}{4\pi} \mathbf{E}_p \times \mathbf{B}_T + \omega \mathbf{S}_L^p = \frac{I}{2\pi} \left(\frac{\omega}{\alpha} \mathbf{B}_p - \frac{1}{\varpi^2} \mathbf{E}_p \times \mathbf{m} \right). \end{aligned} \quad (\text{F.161})$$

Thus, at the neutron star surface where $\alpha = \alpha(r_s) \neq 0$,

$$\begin{aligned} -\mathbf{S}_L \cdot \mathbf{n} &\rightarrow \frac{dJ}{d\Sigma_s dt} = -\frac{I}{2\pi\alpha} B_\perp = -\frac{I}{4\pi^2\alpha\varpi} (\nabla\Psi \times e_\phi) \cdot \mathbf{n} \\ -\mathbf{S}_E \cdot \mathbf{n} &\rightarrow \frac{dM}{d\Sigma_s dt} = -\frac{I}{2\pi} \left[\frac{\omega}{\alpha} B_\perp - \frac{1}{\varpi} (\mathbf{E}_p \times e_\phi) \cdot \mathbf{n} \right] \\ &= -\frac{I}{2\pi\alpha} \Omega_F B_\perp = \Omega_F \frac{dJ}{d\Sigma_s dt} \end{aligned} \quad (\text{F.162})$$

where \mathbf{n} denotes the unit vector outer normal to the neutron star surface. Now note that when the spin J of the rotating neutron star and the magnetic field \mathbf{B} are parallel, $B_\perp > 0$, $I > 0$ whereas when J and \mathbf{B} are antiparallel, $B_\perp < 0$, $I < 0$ due to their definitions (16) and (19). Namely, the magnetic flux (and B) is defined to be positive/negative when it directs *upward/downward* while the poloidal current is defined to be positive/negative when it directs *downward/upward* as we noted

earlier. Thus one always has $IB_{\perp} > 0$, and hence from eq.(41) above, we always have

$$\begin{aligned}-\mathbf{S}_L \cdot \mathbf{n} &= -\frac{I}{2\pi\alpha} B_{\perp} < 0, \\ -\mathbf{S}_E \cdot \mathbf{n} &= -\frac{I}{2\pi\alpha} \Omega_F B_{\perp} < 0.\end{aligned}\quad (\text{F.163})$$

Since the angular momentum and the energy flux going *into* the neutron star surface are all *negative*, this means that the rotating neutron star (i.e. the pulsar) experiences magnetic braking torque, namely spins-down and as a result, always loses part of its rotational energy (at the surface).

9.8 (i) Substituting $\mathbf{E} = \mathbf{E}_p = -(\Omega_F - \omega)/(2\pi\alpha)\nabla\Psi$ into the Maxwell equation $\nabla \times (\alpha\mathbf{E}) = (\mathbf{B} \cdot \nabla\omega)\mathbf{m}$, one can readily realize that

$$(\mathbf{B} \cdot \nabla)\Omega_F = 0 \quad (\text{F.164})$$

indicating that Ω_F is constant on magnetic surfaces, i.e. $\Omega_F = \Omega_F(\Psi)$, which represents the generalized Ferraro's isorotation law.

(ii) Combining

- the *freezing-in condition*: $\mathbf{E}_T + \frac{1}{c}(\mathbf{v} \times \mathbf{B})_T = 0$,
- the *particle conservation*: $\nabla \cdot (\alpha\gamma n\mathbf{v}) = 0$,
- the *Maxwell equation*: $\nabla \cdot \mathbf{B} = 0$

one ends up with $\mathbf{u}_p = \gamma\mathbf{v}_p = \eta(\mathbf{B}_p/\alpha n)$ and hence from

$$\mathbf{u}_T = \gamma\mathbf{v}_T = \eta\left(\frac{1}{\alpha n}\mathbf{B}_T\right) + \gamma\left[\frac{\Omega_F - \omega}{\alpha}\right]\mathbf{e}_{\phi} \quad (\text{F.165})$$

it follows that

$$\mathbf{u} = \gamma\mathbf{v} = \frac{\eta}{\alpha n}\mathbf{B} + \gamma\left[\frac{\Omega_F - \omega}{\alpha}\right]\mathbf{e}_{\phi}, \quad (\text{F.166})$$

where the quantity η represents the *particle flow along the magnetic flux* or the *particle-to-magnetic field flux ratio*.

Then plugging \mathbf{u} back into the particle number conservation yields

$$\begin{aligned}0 &= \nabla \cdot (\alpha n\mathbf{u}) = \nabla \cdot (\eta\mathbf{B}) \\ &= \eta(\nabla \cdot \mathbf{B}) + (\mathbf{B} \cdot \nabla)\eta = (\mathbf{B} \cdot \nabla)\eta,\end{aligned}\quad (\text{F.167})$$

which implies that η must be constant on magnetic surfaces as well, i.e. $\eta = \eta(\Psi)$.

(iii)–(iv)

Let ξ^{μ} be a Killing field associated with an isometry of the background spacetime metric, then $\nabla_{\nu}(T^{\mu\nu}\xi_{\mu}) = 0$. Since stationary and axisymmetric spacetimes have

two Killing fields $k^\mu = (\partial/\partial t)^\mu$ and $m^\mu = (\partial/\partial\phi)^\mu$, respectively, the energy flux and angular momentum flux vector

$$P_E^\mu = -T^{\mu\nu}k_\nu, \quad P_L^\mu = T^{\mu\nu}m_\nu \quad (\text{F.168})$$

are covariantly conserved. Thus using,

$$T_p^{\mu\nu} + T_{\text{em}}^{\mu\nu} = n\mu u^\mu u^\nu + Pg^{\mu\nu} + \frac{1}{4\pi} \left\{ F_\alpha^\mu F^{\nu\alpha} - \frac{1}{4} g^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right\} \quad (\text{F.169})$$

$$u^\mu = (\gamma, \gamma\mathbf{v}) = \left(\gamma, \frac{\eta}{n\alpha} \mathbf{B} + \gamma \left[\frac{\Omega_F - \omega}{\alpha} \right] \mathbf{e}_\phi \right) \quad (\text{F.170})$$

and

$$P_E^A = -T_t^A = nu^A E, \quad (A = r, \theta) \quad (\text{F.171})$$

$$P_L^A = T_\phi^A = nu^A L, \quad (\text{F.172})$$

one gets two more integrals of motion [72, 97, 98]

$$E = E(\Psi) = \frac{\Omega_F I}{2\pi} + \mu\eta(\alpha\gamma + \omega\mathbf{u}_\phi), \quad (\text{F.173})$$

$$L = L(\Psi) = \frac{I}{2\pi} + \mu\eta u_\phi$$

together with the total loss of energy and angular momentum given by

$$W_{\text{tot}} = \int_0^{\Psi_{\text{max}}} E(\Psi) d\Psi, \quad (\text{F.174})$$

$$K_{\text{tot}} = \int_0^{\Psi_{\text{max}}} L(\Psi) d\Psi. \quad (\text{F.175})$$

(v) The entropy conservation $\nabla_\alpha(nsu^\alpha) = 0$ reduces, for stationary axisymmetric case, to

$$\nabla \cdot (\alpha ns\mathbf{u}) = 0. \quad (\text{F.176})$$

Thus using

$$\mathbf{u} = \frac{\eta}{\alpha n} \mathbf{B} + \gamma \left[\frac{\Omega_F - \omega}{\alpha} \right] \mathbf{e}_\phi, \quad (\text{F.177})$$

one gets

$$\begin{aligned} 0 &= \nabla \cdot (\alpha ns\mathbf{u}) = \nabla \cdot (\eta s\mathbf{B}) \\ &= s\nabla \cdot (\eta\mathbf{B}) + \eta(\mathbf{B} \cdot \nabla)s = \eta\mathbf{B} \cdot \nabla)s \end{aligned} \quad (\text{F.178})$$

which implies that the entropy per particle s must be constant on magnetic surfaces as well, $s = s(\Psi)$. To summarize, for the stationary axisymmetric case, there are five-integrals of motion (constants on magnetic surfaces)

$$\{\Omega_F(\Psi), \eta(\Psi), s(\Psi), E(\Psi), L(\Psi)\}. \quad (\text{F.179})$$

We shall now show that once the poloidal magnetic field B_p and the five-integrals of motion given above are known, the toroidal magnetic field B_ϕ and all the other plasma parameters characterizing a plasma flow can be determined. To do so, we solve the two conservation laws and the toroidal component

$$u_\phi = \frac{\eta}{\alpha n} B_\phi + \gamma \left[\frac{\Omega_F - \omega}{\alpha} \right] \varpi = -\frac{2\eta I}{\alpha^2 n \varpi} + \gamma \left[\frac{\Omega_F - \omega}{\alpha} \right] \varpi \quad (\text{F.180})$$

for the Lorentz factor γ , the angular momentum u_ϕ and the poloidal current flux function I to get [72, 98]

$$\gamma(\Psi, r) = \frac{E}{\alpha n \mu} \frac{\alpha^2(1 - \Omega_F L/E) - M^2(1 - \omega L/E)}{\alpha^2 - (\Omega_F - \omega)^2 \varpi^2 - M^2} \quad (\text{F.181})$$

$$u_\phi(\Psi, r) = \frac{E}{\varpi \eta \mu} \frac{(1 - \Omega_F L/E)(\Omega_F - \omega) \varpi^2 - M^2 L/E}{\alpha^2 - (\Omega_F - \omega)^2 \varpi^2 - M^2} \quad (\text{F.182})$$

$$I(\Psi, r) = 2\pi\eta \frac{\alpha^2 L - (\Omega_F - \omega) \varpi^2 (E - \omega L)}{\alpha^2 - (\Omega_F - \omega)^2 \varpi^2 - M^2}. \quad (\text{F.183})$$

The quantity

$$M^2 \equiv \frac{4\pi\mu\eta^2}{n} = \frac{\alpha^2 u_p^2}{u_A^2} \quad (\text{F.184})$$

is the square of the **Mach number** of the poloidal velocity $u_p = \eta B_p / n\alpha$ with respect to the Alfvén velocity $u_A = B_p / \sqrt{4\pi n \mu}$. The above expressions have critical points, where the denominator vanishes, i.e. at positions along the flow, where

$$\alpha^2(r_A) - (\Omega_F - \omega(r_A))^2 \varpi_A^2 = M_A^2. \quad (\text{F.185})$$

These are the Alfvén points along the flow. For given rotation, this equation has two solutions, an inner Alfvén point and an outer one. In order to get regular expressions at the Alfvén points, the nominators also have to vanish. This fixes the total angular momentum L

$$M_A^2 L/E = \varpi_A^2 (\Omega_F - \omega(r_A)) (1 - \Omega_F L/E). \quad (\text{F.186})$$

In order to determine this Mach number, consider

$$\gamma^2 - \mathbf{u}^2 = \gamma^2 - \gamma^2 \mathbf{v}^2 = \gamma^2 (1 - \mathbf{v}^2) = 1 \quad (\text{F.187})$$

and into this relation, we substitute the above expressions to get

$$\frac{F_K}{\varpi^2 D^2} = \frac{1}{64\pi^4} \frac{M^4 (\nabla \Psi)^2}{\varpi^2} + \alpha^2 \eta^2 \mu^2, \quad (\text{F.188})$$

where

$$D = \alpha^2 - (\Omega_F - \omega)^2 \varpi^2 - M^2 \quad (\text{F.189})$$

$$\begin{aligned} F_K &= \alpha^2 \varpi^2 (E - \Omega_F L)^2 [\alpha^2 - (\Omega_F - \omega)^2 \varpi^2 - 2M^2] \\ &\quad + M^4 [\varpi^2 (E - \omega L)^2 - \alpha^2 L^2]. \end{aligned} \quad (\text{F.190})$$

D is the generalization of the light cylinder function defined for force-free magnetospheres. This is the Bernoulli equation. We bring the wind equation into dimensionless form by scaling radii with the asymptotic light cylinder, $R_L = c/\Omega_F$, $x = \varpi/R_L$,

$$\alpha^2 x^2 \frac{F_K(x; M^2, \epsilon)}{D^2(x; M^2)} \left(\frac{E}{\mu} \right)^2 = \alpha^2 x^4 + \frac{B_p^2 x^4}{16\pi^2 \mu^2 \eta^2} M^4. \quad (\text{F.191})$$

The parameter $\epsilon = \Omega_F L/E$ is a measure for the inertia of the plasma, with $\epsilon = 1$ in the force-free limit. The last term on the right-hand side can be scaled to the foot point of the magnetic flux surface

$$\Phi_\Psi^{-1}(x) = \frac{B_p \varpi^2}{B_{p,*} \varpi_*^2}. \quad (\text{F.192})$$

One of the essential parameters which determine the plasma flow along the flux tube is then given by *Michel's magnetization parameter* σ_* defined as follows [97, 98]

$$\sigma_*(\Psi) = \frac{(B_{p,*} \varpi_*^2) c}{4\pi \mu \eta(\Psi) R_L^2(\Psi)}. \quad (\text{F.193})$$

The asymptotic Lorentz factor achieved in the plasma flow along a given flux surface is then essentially determined by the magnetization [98]. Highly relativistic flows are achieved for $\sigma_* \gg 1$.

To summarize, once B_p , $\Omega_F(\Psi)$, $\eta(\Psi)$, $s(\Psi)$, $E(\Psi)$, $L(\Psi)$ are known, the characteristics of the plasma flow, I , γ , u_ϕ , u_p , M^2 can be determined by the above wind equation (F.188). For more details, see [98, 157, 158].

Problems in Chapter 10

10.1 MRI Dispersion Relation

The complete dispersion relation for MRI can be found in [49]. We consider perturbations of the form $\exp(i[\mathbf{k} \cdot \mathbf{x} - \omega t])$. The perturbed MHD equations lead to the following dispersion relation

$$\begin{aligned} &[\omega^2 - (\mathbf{k} \cdot \mathbf{V}_A)^2] [\omega^4 - k^2(c_S^2 + V_A^2)\omega^2 + (\mathbf{k} \cdot \mathbf{V}_A)^2 k^2 c_S^2] \\ &- \left[\kappa^2 \omega^4 - \omega^2 (\kappa^2 k^2 (c_S^2 - V_{A\phi}^2) + \left(\mathbf{k} \cdot \mathbf{V}_A \right)^2 \frac{d\Omega^2}{d \ln R} \right] \\ &- k^2 c_S^2 (\mathbf{k} \cdot \mathbf{V}_A)^2 \frac{d\Omega^2}{d \ln R} = 0. \end{aligned} \quad (\text{F.194})$$

κ is the epicycle frequency defined as

$$\kappa^2 = \frac{d\Omega^2}{d \ln R} + 4\Omega^2. \quad (\text{F.195})$$

This is a third-order equation in ω^2 which has three solution branches. In the non-rotating case, these solutions are the normal MHD waves, the Alfvén waves and the slow and fast magnetosonic waves. For Keplerian rotation, the slow wave becomes unstable for $\Omega^2 = (\mathbf{k} \cdot \mathbf{V}_A)^2/3$. For this, make a plot of ω^2 vs. Ω^2 .

In the Boussinesque approximation, $c_S^2 \rightarrow \infty$, the dispersion relation is simplified to a biquadratic equation

$$\omega^4 - \omega^2 [\kappa^2 + 2(\mathbf{k} \cdot \mathbf{V}_A)^2] + (\mathbf{k} \cdot \mathbf{V}_A)^2 \left((\mathbf{k} \cdot \mathbf{V}_A)^2 + \frac{d\Omega^2}{d \ln R} \right) = 0. \quad (\text{F.196})$$

ω^2 will be negative, if

$$(\mathbf{k} \cdot \mathbf{V}_A)^2 < -\frac{d\Omega^2}{d \ln R}. \quad (\text{F.197})$$

Make a plot of the growth rate in units of Keplerian angular velocity as a function of the wavenumber, $\mathbf{k} \cdot \mathbf{V}_A$.

10.2 Ring Diffusion

The calculation is straightforward.

10.3 Relativistic Keplerian Disks

References for a modern treatment of relativistic Keplerian disks can be found in [253].

10.4 Radiative Transfer around Rotating Black Holes

(i) The metric tensor of the Kerr geometry in cylindrical coordinates is given by

$$g_{tt} = -1 + \frac{2M}{R} - \frac{Mz^2}{R^3} \left(1 + \frac{2a^2}{R^2} \right) \quad (\text{F.198})$$

$$g_{t\phi} = -\frac{2aM}{r} + \frac{aMz^2}{r^3} \left(3 + \frac{2a^2}{r^2} \right) \quad (\text{F.199})$$

$$g_{\phi\phi} = r^2 + a^2 + \frac{2Ma^2}{r} - \frac{a^2z^2}{r^2} \left(1 + \frac{5M}{r} + \frac{2Ma^2}{r^3} \right) \quad (\text{F.200})$$

$$g_{RR} = \frac{1}{\mathcal{A}} - \frac{z^2}{r^2 \mathcal{A}^2} \left[\frac{M}{r} \left(3 - \frac{4M}{r} \right) - \frac{a^2}{r^2} \left(3 - \frac{6M}{r} + \frac{2a^2}{r^2} \right) \right] \quad (\text{F.201})$$

$$g_{Rz} = \frac{z}{R\mathcal{A}} \left(\frac{2M}{R} - \frac{a^2}{R^2} \right) \quad (\text{F.202})$$

$$g_{zz} = 1 + \frac{z^2}{R^2 \mathcal{A}} \left(\frac{2M}{R} - \frac{2Ma^2}{R^3} + \frac{a^4}{R^4} \right). \quad (\text{F.203})$$

(ii) The dominant part of the radial Euler equation leads to

$$(U^t)^2 \Gamma_{tt}^2 + 2U^t U^\phi \Gamma_{t\phi}^2 + (U^R)^2 \Gamma_{\phi\phi}^2 = 0. \quad (\text{F.204})$$

Together with the normalization condition $U^\mu U_\mu = -1$, this yields the desired results for U^t and U^ϕ .

(iii) The transformation matrices L are explicitly given by

$$L = \begin{pmatrix} -U_t & -U_\phi & 0 & 0 \\ L_1^t & L_\phi^\phi & 0 & 0 \\ 0 & 0 & 1/\sqrt{\mathcal{A}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{F.205})$$

and the inverse relation

$$\bar{L} = \begin{pmatrix} U^t & \bar{L}^t & 0 & 0 \\ U^\phi & \bar{L}_\phi^\phi & 0 & 0 \\ 0 & 0 & \sqrt{\mathcal{A}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\text{F.206})$$

where

$$\begin{aligned} L_t^\phi &= -\sqrt{\frac{\mathcal{A}}{\mathcal{B}}} \sqrt{\frac{M}{R}} & L_\phi^\phi &= R \sqrt{\frac{\mathcal{A}}{\mathcal{B}}} \left(1 + \frac{a}{R} \sqrt{\frac{M}{R}}\right) \\ \bar{L}_t^\phi &= \frac{1}{\sqrt{\mathcal{A}\mathcal{B}}} \left(1 + \frac{a^2}{R^2} - \frac{2a}{R} \sqrt{\frac{M}{R}}\right), & \bar{L}_\phi^\phi &= \frac{1}{R\sqrt{\mathcal{A}\mathcal{B}}} \left(1 - \frac{2M}{R} + \frac{a}{R} \sqrt{\frac{M}{R}}\right) \end{aligned} \quad (\text{F.207})$$

(iv) The vector m^i is given by

$$m^i = v^2 \left[\omega_{tt}^i + (\omega_{tk}^i + \omega_{kt}^i) n^k + \omega_{jk}^i n^j n^k \right] \quad (\text{F.208})$$

with the expression for transfer equation

$$\begin{aligned} m^i \frac{\partial f}{\partial \bar{p}^i} &= m^t \frac{\partial f}{\partial v} + \left[m^t \sqrt{1-\mu^2} - m^1 \sin \chi - m^2 \cos \chi \right] \frac{\sqrt{1-\mu^2}}{\mu v} \frac{\partial f}{\partial \mu} \\ &\quad + \frac{-m^1 \cos \chi + m^2 \sin \chi}{v \sqrt{1-\mu^2}} \frac{\partial f}{\partial \chi}. \end{aligned} \quad (\text{F.209})$$

(v) The lowest order in the transfer equation has the expression

$$\begin{aligned} &v\sqrt{\mathcal{A}}\sqrt{1-\mu^2} \cos \chi \frac{\partial f}{\partial R} + v\mu \frac{\partial f}{\partial z} + \frac{v}{r\sqrt{\mathcal{B}}} \left[\sqrt{\frac{M}{r}} + \frac{\mathcal{E}}{\sqrt{\mathcal{A}}} \sqrt{1-\mu^2} \sin \chi \right] \frac{\partial f}{\partial \phi} \\ &+ \frac{3\mathcal{A}}{2\mathcal{B}} \sqrt{\frac{M}{r^3}} (1-\mu^2) \sin \chi \cos \chi v^2 \frac{\partial f}{\partial v} \\ &- \left[\frac{3\mathcal{A}}{2\mathcal{B}} \sqrt{\frac{M}{r^3}} (1-\mu^2) \sin \chi + \frac{\mathcal{A}-1}{r\sqrt{\mathcal{A}}} \right] v\mu \sqrt{1-\mu^2} \cos \chi \frac{\partial f}{\partial \mu} \\ &+ \left[\frac{3\mathcal{A}}{2\mathcal{B}} \sqrt{\frac{M}{r^3}} \cos^2 \chi - \frac{\sin \chi}{r\sqrt{\mathcal{A}}\sqrt{1-\mu^2}} \left(1 - \frac{M}{r} - \mu^2 \mathcal{F}\right) - 2\sqrt{\frac{M}{r^3}} \right] v \frac{\partial f}{\partial \chi} = \bar{\mathcal{Q}}, \end{aligned} \quad (\text{F.210})$$

where

$$\mathcal{E} = 1 - \frac{2M}{r} + \frac{a}{r} \sqrt{\frac{M}{r}}, \quad \mathcal{F} = 1 + \frac{M}{r} - \frac{a^2}{r^2}. \quad (\text{F.211})$$

For an axisymmetric radiation field, we have

$$\frac{\partial f}{\partial \phi} = 0, \quad \int_0^{2\pi} \sin \chi \, f \, d\chi = 0. \quad (\text{F.212})$$

After integration of the transfer equation over χ , we obtain the simple equation

$$\mu v^3 \frac{\partial f}{\partial z} = \mu \frac{\partial \bar{I}_v}{\partial z} = v^2 \bar{Q}, \quad (\text{F.213})$$

where $\bar{I}_v = \bar{I}_v(R, z, \mu)$ is the specific intensity as measured in the local plasma frame LRFM.

(viii) The relevant interactions between plasma and photons near the horizon of a black hole are Bremsstrahlung and Compton scattering, which gives the scattering operator in the plasma frame

$$\begin{aligned} v^2 \bar{Q} &= \kappa_{ff} \varrho_0 [B_v - \bar{I}_v] \\ &+ \kappa_T \varrho_0 \frac{v}{m_e} \frac{\partial}{\partial v} \left[v T_e B_v \frac{\partial}{\partial v} \left(\frac{\bar{I}_v}{B_v} \right) + \frac{\bar{I}_v}{2v^2} (\bar{I}_v - B_v) \right]. \end{aligned} \quad (\text{F.214})$$

κ_{ff} is the Bremsstrahlung opacity and $\kappa_T = 0.4 \text{ cm}^2/\text{g}$ the Thomson opacity. The differential operator describes Compton scattering in the Fokker–Planck approximation [13].

10.6 The solution of this problem is discussed in [36].

Glossary

Accretion disk: Flat disk of matter spiraling down onto the surface of a star or into a black hole. Often, the matter originated on the surface of a companion star in a binary system. Viscosity within the disk generates heat and saps orbital momentum, causing material in the disk to spiral inward, until it impacts in an accretion shock on the central body if the body is a star, or slips toward the event horizon if the central body is a black hole. The most spectacular accretion disks found in nature are those of active galactic nuclei and quasars, which are believed to be supermassive black holes at the center of galaxies. The accretion disk of a black hole is hot enough to emit X-rays just outside of the event horizon. In the modern view, an accretion disk is a quasistationary solution of radiative magnetohydrodynamics, provided the initial configuration has sufficient gas, angular momentum and magnetic fields.

Active galactic nucleus (AGN): The central region of a galaxy that shows unusual energetic activity.

Angular resolution: The ability of a telescope to distinguish two adjacent objects on the sky, or to study the fine details on the surface of some object; often synonymous with “clarity” or “sharpness.”

Arcsecond (arcsec): A unit of angular measure of which there are 60 in 1 arcminute (or therefore 3600 in 1 arc degree).

Binary star system: A system which consists of two stars orbiting about their common center-of-mass, held together by their mutual gravitational attraction. Most stars are found in binary star systems.

Black hole: A dense, compact object whose gravitational pull is so strong that, within a certain distance of it, nothing can escape, not even light. Black holes are thought to result from the collapse of certain very massive stars at the ends of their evolution. Current theories predict that all the matter in a black hole is piled up in a single point (or ring, when rotating) at the center, but we do not understand how this central singularity works. To properly understand the black hole center requires a fusion of the theory of gravity with the theory that describes the behavior of matter

on the smallest scales, called quantum mechanics. This unifying theory has already been given a name, quantum gravity, but how it works is still unknown.

Blazars: A class of active galaxies that exhibit rapidly variable emission from the radio through gamma-ray band. The radiation is predominantly from jets moving near the speed of light. Blazars are thought to be radio galaxies with their jets oriented toward Earth.

Cataclysmic system: Cataclysmic variables are a class of binary stars containing a white dwarf and a companion star. The companion star is usually a red dwarf, although in some cases it is another white dwarf or a slightly evolved star (subgiant). Several hundreds of cataclysmic variables are known. The stars are so close to each other that gravity of the white dwarf distorts the secondary, and the white dwarf accretes matter from the companion.

Cauchy horizon: A Cauchy horizon is a light-like boundary of the domain of validity of a Cauchy problem (a boundary value problem of the theory of partial differential equations). One side of the horizon contains closed space-like geodesics and the other side contains closed time-like geodesics. The simplest example is the internal horizon of a Reissner–Nordström black hole. It also appears in the Kerr black hole.

CFL condition: The Courant–Friedrichs–Levy condition, usually abbreviated to the CFL condition, says that in any time-marching computer simulation the time-step must be less than the time for some significant action to occur, and preferably considerably less. For example, if we have a computer simulation of a satellite orbiting a planet, the discrete time interval which we use must obviously be less than the orbital period on common sense grounds. For the sake of stability, it must be less than one quarter of the orbital period, and, in practice, one will take a step of about one fortieth of the orbital period. The CFL condition was originally formulated in the context of compressible fluid flows. If we divide the flow volume up into cells, then we need a time-step less than the time taken for a sound wave to cross one of the cells.

Chandrasekhar mass: The upper limit to the mass of a white dwarf, equals $(5.87/\mu^2)M_\odot$, where μ is the mean number of nucleons per electron. For Fe we have $\mu = 56/26$.

Chandra X-ray Observatory (CXO): Formerly called AXAF, Chandra was launched July 23, 1999, and is with XMM–Newton the largest and most sophisticated X-ray observatory to date. NASA’s Chandra X-ray Observatory was named in honor of the late Indian–American Nobel laureate, Subrahmanyan Chandrasekhar. The Chandra spacecraft carries a high-resolution mirror, two imaging detectors, and two sets of transmission gratings. Important Chandra features are: an order of magnitude improvement in spatial resolution, and good sensitivity from 0.1 to 10 keV.

Color superconductivity: Color superconductivity is a phenomenon predicted to occur in quark matter if the baryon density is sufficiently high (well above nuclear density) and the temperature is not too high (i.e. below 10^{12} kelvin). Color-superconducting phases are to be contrasted with the normal phase of quark matter, which is just a weakly interacting Fermi liquid of quarks. Unlike an electrical superconductor, color-superconducting quark matter comes in many varieties, each of which is a separate phase of matter. In forming the Cooper pairs, there is a 9×9 color-flavor matrix of possible pairing patterns. The differences between these patterns are very physically significant: different patterns break different symmetries of the underlying theory, leading to different excitation spectra and different transport properties. In theoretical terms, a color-superconducting phase is a state in which the quarks near the Fermi surface become correlated in Cooper pairs, which condense. In phenomenological terms, a color-superconducting phase breaks some of the symmetries of the underlying theory, and has a very different spectrum of excitations and very different transport properties from the normal phase.

Compton Gamma Ray Observatory (CGRO): The Compton Gamma Ray Observatory was the second of NASA's Great Observatories. CGRO, at 17 tonnes, was the heaviest astrophysical payload ever flown at the time of its launch on April 5, 1991 aboard the space shuttle Atlantis. Compton was safely deorbited and re-entered the Earth's atmosphere on June 4, 2000. CGRO had four instruments that covered six decades of the electromagnetic spectrum, from 30 keV to 30 GeV. In order of increasing spectral energy coverage, these instruments were the Burst And Transient Source Experiment (BATSE), the Oriented Scintillation Spectrometer Experiment (OSSE), the Imaging Compton Telescope (COMPTEL), and the Energetic Gamma Ray Experiment Telescope (EGRET).

Compton scattering: The scattering, or collision, of a photon with an electron.

Comptonization: The X-ray power-law component in galactic black hole candidates and AGN is attributed to black-body photons being up-scattered by electrons (that is, the photons gain energy from the electrons via the inverse Compton process), as they traverse a hot plasma with a Maxwellian electron temperature $10\text{--}100$ keV. Comptonization is saturated, where the photons are in thermal equilibrium with the electrons. Then a cutoff in the spectrum occurs at $\sim 3k_B T_e$, where T_e is the temperature of the hot electrons.

Continuous spectrum: Spectrum in which the radiation is distributed over all frequencies, not just a few specific frequency ranges. A prime example is the black-body radiation emitted by a hot, dense body.

Core: The central region of a planet, star, neutron star, or a galaxy.

Corona: The outermost atmosphere of a star (including the Sun) or of an accretion disk, millions of kilometers in extent, and consisting of highly rarefied gas heated to temperatures of millions of degrees.

Cosmic radiation: Very energetic radiation from outer space which hits the Earth's atmosphere and can still be detected near the surface after various transformations. It consists mostly of highly energetic particles (protons, helium cores, heavy atomic nuclei and leptons) and X-rays and gamma radiation. Only a small fraction of cosmic radiation is produced in the Sun, the rest has its origin in partly still unknown sources inside and outside of the Milky Way.

Cosmological constant: A modification of the equations of general relativity that represents a possible repulsive force in the Universe. The cosmological constant could be due to the energy density of the vacuum.

Cygnus A: This galaxy is the brightest radio source (as indicated by the letter A) in the constellation Cygnus (Swan). The supermassive black hole in its center is a billion times heavier than the Sun. Although the galaxy is relatively distant (300 times further away than the Andromeda Galaxy), it appears to us as the second brightest radio source in the entire sky. This is because the black hole generates tremendous energy as it consumes large amounts of material. Nearby electrons are accelerated in this process, emitting strong radio waves as they spiral outward in magnetic fields.

Cygnus X-1: This is the brightest X-ray source (indicated as X-1) in the constellation Cygnus (Swan). It consists of a bright blue star and a black hole that orbit around each other. The black hole pulls gas off the surface of this star. This gas heats up and shines in X-rays as it falls towards the black hole.

Density: A measure of the compactness of the matter within an object, computed by dividing the mass by the volume of the object.

Eccentricity: A measure of the flatness of an ellipse, equal to the distance between the two foci divided by the length of the major axis.

Eddington luminosity: The limit beyond which the radiation force on matter is greater than the gravitational force. This limit is independent on the radius of the emitting surface. The corresponding luminosity is $L_{\text{Ed}} = 4\pi G M m_p c / \sigma_T$ is proportional to the mass M of the object.

Electron-volt (eV): The energy gained by an electron accelerated by a potential of 1 volt. One electron-volt corresponds to a frequency $\nu = 2.418 \times 10^{14}$ Hz in electromagnetic radiation, or a temperature of 11.606 K.

Ergosphere: The region of a rotating Kerr black hole between the static surface and the event horizon. In this region, everything is forced to rotate in the same sense as the black hole, although you can still escape.

Escape velocity: The speed necessary for an object to escape the gravitational pull of an object. Anything that moves away from the object with more than the escape velocity will never return.

Event horizon: Imaginary spherical surface surrounding a black hole, with radius equal to the Schwarzschild radius, within which no event can be seen heard, or known about by an outside observer.

Excited state: The state of an atom when one of its electrons is in a higher energy orbital than the ground state. Atoms can become excited by absorbing a photon of a specific energy, or by colliding with a nearby atom.

Fluorescence: The absorption of a photon of one energy, or wavelength, and re-emission of one or more photons at lower energies, or longer wavelengths.

Gamma-ray: Region of the electromagnetic spectrum, beyond X-rays, corresponding to radiation of very high frequency and very short wavelength.

Gamma-ray burst (GRB): An outburst that radiates tremendous amounts of energy equal to or greater than a supernova, in the form of gamma-rays and X-rays, with a duration from a few milliseconds to thousands of seconds. GRBs are isotropically distributed on the Sky.

Gamma-Ray Large Area Space Telescope (GLAST): GLAST is a next generation high-energy gamma-ray observatory designed for making observations of celestial gamma-ray sources in the energy band extending from 10 MeV to more than 100 GeV to be launched in 2007. It follows in the footsteps of the CGRO-EGRET experiment, which was operational between 1991 and 1999. The GLAST LAT has a field of view about twice as wide (more than 2.5 steradians), and sensitivity about 50 times that of EGRET at 100 MeV and even more at higher energies.

Globular cluster: Tightly bound, roughly spherical collection of a few hundred thousand of stars spanning about 100 lightyears. Globular clusters are distributed in the haloes around the Milky Way and other galaxies.

Gluons: Term for exchange particles of the strong interaction (derived from glue). There are eight different gluons transmitting the force between the quarks. They are electrically neutral and massless.

Gravitational instability: A condition whereby an object's (inward-pulling) gravitational potential energy exceeds its (outward-pushing) thermal energy, thus causing the object to collapse.

Gravitational lensing: Banding of light from a distant object by a massive foreground object (a star, a galaxy, or a cluster of galaxies).

Gravitational redshift: A prediction of Einstein's general theory of relativity. Photons lose energy as they escape the gravitational field of a compact object. Because a photon's energy is proportional to its frequency, a photon that loses energy suffers a decrease in frequency, or redshift, in wavelength.

Gravitational wave: The gravitational analog of an electromagnetic wave, whereby gravitational radiation is emitted at the speed of light from any mass that undergoes rapid acceleration.

GRS 1915+105: GRS 1915+105 is a microquasar, a galactic object that has been associated with relativistic jets and extremely variable radio, infrared, and X-ray emission. It is thought to be a binary system containing a black hole that is accreting matter from a stellar companion.

Hertzsprung–Russell (HR) diagram: A plot of luminosity vs. temperature for a group of stars that can be used to classify the evolutionary state of stars.

Horizontal branch: Region of the HR diagram where post-main-sequence stars again reach hydrostatic equilibrium. At this point, the star is burning helium in its core, and hydrogen in a shell surrounding the core.

Hybrid (neutron) stars: Neutron stars consisting of normal matter in the outer parts and a quark-matter core. Quark matter is probably in a color-superconducting state (2SC or quark-flavor locked (CFL) phase).

INTEGRAL: INTEGRAL (INTERnational Gamma-Ray Astrophysics Laboratory) is an astronomical satellite for observing the gamma-ray sky. It was selected by the ESA (European Space Agency) science program committee on June 3rd 1993 as a medium size mission. The INTEGRAL satellite was launched on October 17, 2002 by a Russian PROTON launcher. It has a highly eccentric orbit with a revolution period around the Earth of three sidereal days. The perigee is at 10,000 km and the apogee at 152,600 km with an inclination of 51.6 degrees with respect to the equatorial plane. The INTEGRAL science payload consists of two main instruments, the spectrometer SPI and the imager IBIS supplemented by two subsidiary instruments, the X-ray monitor JEM-X and the optical monitoring camera OMC.

Interferometry: Technique in widespread use to dramatically improve the resolution of telescopes, especially radio telescopes. Several radio telescopes observe the object simultaneously, and a computer analyzes how the signals interfere with each other.

Ionization: The process by which ions are produced, typically by collisions of electrons, ions, or photons.

Innermost stable circular orbit (ISCO): This radius marks the location of the innermost stable circular orbit around a black hole. Outside three Schwarzschild radii, all circular orbits are stable, meaning that a small blast on the manoeuvering thrusters by a rocket in circular orbit would not perturb the orbit greatly.

Inverse Compton emission: In physics, Compton scattering or the Compton effect, is the decrease in energy (increase in wavelength) of an X-ray or gamma-ray photon, when it interacts with matter. Inverse Compton scattering indicates the effect, where the photon gains energy (decreasing in wavelength) upon interaction with matter. Inverse Compton scattering is important in astrophysics. In X-ray astronomy, the accretion disk surrounding a black hole is believed to produce a thermal spectrum. The lower energy photons produced from this spectrum are scattered to higher energies by relativistic electrons in the surrounding corona. This is the origin of the power-law component in the X-ray spectra (0.2–100 keV) of accreting black holes. The inverse Compton effect is also important in jets, where relativistic electrons scatter low-frequency photons to gamma-rays.

Jet: A highly directed flow of gas or plasma that comes from such a flow.

Kepler's Laws of motion: Three laws which summarize the motions of the planets about the Sun, or more generally, the motion of one star (neutron star, black hole) around another under the influence of gravity.

Kerr black hole: An exact solution of Einstein's field equations that is the metric outside a spinning event horizon found by Roy Kerr in 1963. It is not the solution for a spinning neutron star.

Killing field: In differential geometry, a Killing vector field is a vector field on a Riemannian manifold that preserves the metric. Killing fields are the infinitesimal generators of isometries; that is, flows generated by Killing fields are continuous isometries of the manifold. Killing fields are named for Wilhelm Killing. A Killing vector field satisfies the Killing equation $L_X g = 0$, where L_X is the Lie derivative along X and g is the Riemannian metric on the manifold.

Kiloelectron-volt (keV): A unit used to describe the energy of X-rays, equal to a thousand electron-volts. One kiloelectron-volt corresponds to a frequency $\nu = 2.418 \times 10^{17}$ Hz in X-rays.

Lagrange point: One of five special points in the plane of two massive bodies orbiting one another, where a third body of negligible mass can remain in equilibrium.

Leptons: One of the two groups of matter particles. There are three pairs of leptons, containing each an electrically charged particle and a neutrino: electron and electron neutrino, muon and muon neutrino, tau and tau neutrino. The leptons are influenced by the electromagnetic and the weak interaction.

LHC: Large Hadron Collider, proton collider at CERN (Geneva), which is built using the old LEP tunnel until 2007.

Light-curve: The variation in brightness of a star with time.

Light deflection: The angle by which a light ray is curved by the gravitational field of a massive body. General relativity gives a value twice as large as that which Newtonian physics would provide, assuming that photons have nonzero mass.

Lighthouse model: The leading explanation for pulsars. A small region of the neutron star, near one of the magnetic poles, emits a steady stream of radiation which sweeps past Earth each time the star rotates. Thus the period of the pulses is just the star's rotation period.

Luminosity: One of the basic properties used to characterize stars. Luminosity is defined as the total energy radiated by a star each second, at all wavelengths.

Magnetosphere: A zone of charged particles trapped by a planet's or star's magnetic field (neutron star or black hole), lying above the atmosphere.

Magnetar: A magnetar is a neutron star with an extremely strong magnetic field, typically a thousand times stronger than in a normal neutron star. Its decay powers the emission of copious amounts of high-energy electromagnetic radiation, particularly X-rays and gamma-rays.

Magnetorotational instability (MRI): Accretion disks are stable to hydrodynamic perturbations, and the fluid flow is expected to be laminar. For there to be turbulence, as required for the standard disk model (α disk), this implies that there is some form of nonlinear hydrodynamic instability, or angular momentum transport is due to some other mechanism. Balbus and Hawley proposed in 1991 a mechanism which involves magnetic fields to generate the turbulence, now called magnetorotational instability (MRI). Magnetohydrodynamics is subtly different from that of hydrodynamics. A weak magnetic field acts like a spring. If there is a weak radial magnetic field in an accretion disk, then two gas volume elements will experience a force acting on them. The inner element will have a force acting to slow it down. This causes it to lose energy and angular momentum and move inwards, where due to orbital mechanics it speeds up. The reverse happens to the outer gas element, which moves outwards and slows down. As a consequence, the magnetic field spring is stretched, transferring angular momentum in the process.

Main sequence: A well-defined band on an HR diagram on which most stars tend to be found, running from the top left of the diagram to the bottom right.

Mass-radius relation: The dependence of the radius of a main-sequence star on its mass. The radius rises roughly in proportion to the mass.

Microquasar: Microquasars are stellar mass black holes, that display characteristics of the supermassive black holes found at the centers of some galaxies. For instance, they have radio jets. In the Spring of 1994, Felix Mirabel from Saclay, France, and Luis Rodriguez, from the National Autonomous University in Mexico City, were observing an X-ray-emitting object called GRS 1915+105 (about 40,000 lightyears away). Their time series of VLA observations showed that a pair of objects ejected from GRS 1915+105 were moving apart at an apparently superluminal speed. This was the first time that superluminal motion had been detected in our own Galaxy.

Millisecond pulsar: A pulsar whose period indicates that the neutron star is rotating nearly 1000 times each second.

Neutrino oscillations: Possible solution to the solar neutrino problem, in which the neutrino has a very tiny mass. In this case, the correct number of neutrinos can be produced in the solar core, but on their way to Earth, some can oscillate, or become transformed into other particles, and thus go undetected.

Neutron star: A dense ball of neutrons that remains after a supernova has destroyed the rest of the star. Typically neutron stars are about 20 km across, and contain more mass than the Sun.

Nonthermal radiation: Radiation released by virtue of a fast-moving charged particle (such as an electron) interacting with a magnetic force field or other particles; this process has nothing to do with heat.

Nova: A star that suddenly increases in brightness, often by a factor of as much as 10,000 then slowly fades back to its original luminosity. A nova is the result of an explosion on the surface of a white dwarf star, caused by matter falling onto its surface from the atmosphere of a binary companion.

Nuclear force: The force that binds protons and neutrons within atomic nuclei, and which is effective only at distances less than about 10^{-13} centimeter.

Nucleon: Building block of atoms, i.e. a proton or a neutron.

Nucleus: Dense, central region of an atom, containing both protons and neutrons, and orbited by one or more electrons.

Opacity: A quantity that measures a material's ability to block electromagnetic radiation. Opacity is the opposite to transparency.

Parallax: The apparent motion of a relatively close object with respect to a more distant background as the location of the observer changes.

Parsec: The distance at which a star must lie in order that its measured parallax due to the Earth's orbit around the Sun is exactly 1 arcsecond, equal to 3.3 lightyears.

Period–luminosity relation: A relation between the pulsation period of a Cepheid variable and its absolute brightness. Measurement of the pulsation period allows the distance of the star to be determined.

Planetary nebula: The ejected envelope of a red giant star, spread over a volume roughly the size of our Solar System, with a hot central star that is in the process of becoming a white dwarf star.

Primordial black holes: A primordial black hole is a hypothetical type of black hole that is formed not by the gravitational collapse of a star, but by the extreme density of matter present during the Universe's early expansion.

Proper motion: The angular movement of a star across the sky, as seen from the Earth, measured in seconds of arc per year. This movement is a result of the star's actual motion through space.

Pulsar: Object that emits radiation in the form of rapid pulses with a characteristic pulse period and duration. Generally used to describe the pulsed radiation from a rotating neutron star.

Quality factor: Quantity characterizing the resonance properties of a resonant system, for example a resonant circuit or a cavity resonator. It depends on the average energy of the system and its dissipative power. The higher the quality is, the more focused is the resonance curve. The bandwidth is correspondingly smaller.

Quantum chromodynamics (QCD): Quantum chromodynamics, the gauge theory describing the color strong interaction.

Quarks: A fractionally charged, basic building block of protons, neutrons, and other elementary particles. There are six different quarks. Similar to the leptons, they form a particle group consisting of three particle pairs: up and down quarks, charm and strange quarks and top and bottom quarks. In nature, quarks can occur in pairs (quark and antiquark, known as mesons) or as a three-piece combination of either quarks or antiquarks.

Quasars: Originally, a distant, highly luminous object that looks like a star. Strong evidence now exists that a quasar is produced by gas falling into a supermassive black hole in the center of a galaxy.

Quasinormal modes: Quasinormal modes (QNM) are the modes of energy dissipation of a perturbed object (neutron star or black hole). In this context, a quasinormal mode is a formal solution of linearized differential equations (such as the linearized equations of general relativity constraining perturbations around a black hole solution) with a complex eigenvalue (or frequency). Black holes have many quasi-normal modes (also called ringing modes) that describe the exponential decrease

of asymmetry of the black hole in time, as it evolves towards the perfect spherical shape.

Quasiperiodic Oscillations (QPO's): Variations in the intensity of X-radiation from sources that show periodic behavior for short time intervals, and a variety of periods.

Radio galaxy: Type of active galaxy that emits most of its energy in the form of long-wavelength radiation.

Radio lobe: Roundish region of radio-emitting gas, lying well beyond the center of a radio galaxy.

Red-giant branch: The section of the evolutionary track of a star that corresponds to continued heating from rapid hydrogen shell burning, which drives a steady expansion and cooling of the outer envelope of the star. As the star gets larger in radius and its surface temperature cools, it becomes a red giant.

Redshift: Change in the wavelength of light emitted from a source moving away from us. The relative recessional motion causes the wave to have an observed wavelength longer (and hence redder) than it would if it were not moving. The cosmological redshift is caused by the stretching of space as the Universe expands.

Relativity, general theory: The theory of gravity formulated by Einstein that describes how a gravitational field can be replaced by a curvature of spacetime.

Resolution limit: Measure for the smallest intervals a detector can resolve separately. These can be time intervals (time resolution), differences in energy or wavelength (energy resolution) or spatial distances (spatial resolution).

RHIC: The Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory is a world-class scientific research facility that began operation in 2000, following 10 years of development and construction. RHIC drives two intersecting beams of gold ions head-on, in a subatomic collision. At extremely high energy densities, QCD predicts a new form of matter, consisting of an extended volume of interacting quarks, antiquarks, and gluons. This is the quark–gluon plasma (QGP).

Riemann curvature: In differential geometry, the Riemann curvature tensor is the most standard way to express curvature of Riemannian manifolds, or more generally, any manifold with an affine connection, including torsion. This curvature tensor measures noncommutativity of the covariant derivative. It satisfies several symmetries known as Bianchi identities.

Riemann problem: The Riemann problem is the simplest possible initial value problem for hyperbolic systems. In one spatial dimension, the Riemann (or shock-tube) problem is composed of two uniform states in the infinite domain separated

by a discontinuity at the origin. For the Euler equations, the exact solution of the Riemann problem is well known, self-similar, and consists of a combination of three wave types: shocks, rarefaction waves, and contact discontinuities. Apart from being an important test-bench, the Riemann problem is a basic building block for a large class of modern numerical methods, called upwind or Godunov schemes.

Roche limit: Often called the tidal stability limit, the Roche limit gives the distance from a planet at which the tidal force, due to the planet, between adjacent objects exceeds their mutual attraction. Objects within this limit are unlikely to accumulate into larger objects. The rings of Saturn occupy the region within Saturn's Roche limit.

Roche lobe: An imaginary surface around a star. Each star in a binary system can be pictured as being surrounded by a tear-shaped zone of gravitational influence, the Roche lobe. Any material within the Roche lobe of a star can be considered to be part of that star. During evolution, one member of the binary star can expand so that it overflows its own Roche lobe, and begins to transfer matter onto the other star.

Rossi X-Ray Timing Explorer (RXTE): The Rossi X-ray Timing Explorer (RXTE) was launched on December 30, 1995. RXTE features unprecedented time resolution in combination with moderate spectral resolution to explore the variability of X-ray sources. Time-scales from microseconds to months are covered in an instantaneous spectral range from 2 to 250 keV.

Schwarzschild radius: The distance from the center of a nonrotating black hole such that, if all the mass were compressed within that region, the escape velocity would equal the speed of light. Once a stellar remnant collapses within this radius, light cannot escape and the object is no longer visible. See event horizon.

Shock wave: A wave front marked by an abrupt change in pressure caused by an object or material moving faster than the speed of sound. For example, a sonic boom produced by an aircraft going faster than the speed of sound.

Shapiro time delay: The Shapiro time-delay effect, or gravitational time-delay effect, is one of the four classic Solar System tests of general relativity. The time-delay effect was first noticed in 1964 by Irwin I. Shapiro. Radar signals passing near a massive object takes slightly longer to travel to a target and longer to return (as measured by the observer) than it would if the mass of the object were not present. This also affects the propagation of radio signals emitted by a pulsar in orbit around another star. This allows us to measure the mass of the partner star and the orbital inclination.

Singularity: A point in the Universe where the density of matter and the gravitational field are infinite, such as the center of a black hole.

Spacetime: A synthesis of the three dimensions of space and of a fourth dimension, time; a hallmark of relativity theory.

Spectral class: Classification scheme, based on the strength of stellar spectral lines, which is an indication of the temperature of a star.

Spectroscopic binary: A binary star system which from Earth appears as a single star, but is known to contain more than one star because of the back-and-forth Doppler shifts that are observed as the two stars orbit one another.

Spitzer Space Telescope: NASA's Great Observatory for infrared astronomy was launched in August 2003. Formerly named SIRTF (Space Infrared Telescope Facility), it was renamed in honor of Lyman Spitzer, Jr.

Static limit: The outer boundary of the region around a spinning black hole that is called the ergosphere.

Stellar-mass black hole: A black hole that formed when a massive star died in a supernova explosion and is somewhat more massive than our Sun.

Superconductivity: Property of certain metals, or neutron star matter, at low temperatures. The electrical resistance of the conductor vanishes, so that the electrical current flows without loss. In modern accelerators, often superconducting magnets and high frequency resonators are used which are operated at temperatures near absolute zero. At low temperature, many metals become superconductors. A metal can be viewed as a Fermi liquid of electrons, and below a critical temperature, an attractive phonon-mediated interaction between the electrons near the Fermi surface causes them to pair up and form a condensate of Cooper pairs, which via the Anderson–Higgs mechanism makes the photon massive, leading to the characteristic behavior of a superconductor: infinite conductivity and the exclusion of magnetic fields (Meissner effect).

Superfluidity: Fermionic condensates are a type of superfluid. As the name suggests, a superfluid possesses fluid properties similar to those possessed by ordinary liquids and gases, such as the lack of a definite shape and the ability to flow in response to applied forces. However, superfluids possess some properties that do not appear in ordinary matter. For instance, they can flow at low velocities without dissipating any energy (i.e. zero viscosity). At higher velocities, energy is dissipated by the formation of quantized vortices, which act as holes in the medium where superfluidity breaks down.

Supergravity: In theoretical physics, a supergravity theory is a field theory combining supersymmetry and general relativity. A supergravity theory contains a spin-2 field whose quantum is the graviton. Supersymmetry requires the graviton field to

have a superpartner. This field has spin 3/2 and its quantum is the gravitino. Supergravity theories are often said to be the only consistent theories of interacting massless spin 3/2 fields.

Supermassive black hole: A black hole with a mass much greater than the most massive stars (100 solar masses). The central regions of virtually every galaxy are thought to contain a supermassive black hole of a million solar masses or more. Our Milky Way harbors in its center a supermassive black hole (Sag A*) with 3.5 million solar masses.

Supernova: Explosive death of a star, caused by the sudden onset of nuclear burning (type I), or gravitational collapse followed by an enormously energetic shock wave (type II). One of the most energetic events of the Universe, a supernova may temporarily outshine the rest of the galaxy in which it resides. Supernovae of type Ia (exploding white dwarfs) are cosmic standard candles used to measure the expansion law of the Universe.

Supernova remnant: The expanding glowing remains from a supernova explosion. The Cygnus Loop is an example of a shell-type remnant. As the shock wave from the supernova explosion plows through space, it heats and stirs up any interstellar material it encounters, thus producing a big shell of hot material in space. Plerions resemble the Crab Nebula. These SNRs are similar to shell-type remnants, except that they contain a pulsar in the middle that blows out electron–positron winds.

Supersymmetry (SUSY): One of the most promising candidates for a theory which goes beyond the Standard Model. To every particle, a supersymmetric partner is assigned – an exchange particle for every matter particle and vice versa. Until now, none of these supersymmetric partner particles was detected, so that no experimental proof for the theory of supersymmetry exists yet.

Swift: The Swift Gamma-Ray Burst Explorer carries three instruments to enable the most detailed observations of gamma-ray bursts (GRBs) to date. It carries three coaligned instruments known as the BAT, the XRT, and the UVOT. The XRT and UVOT are X-ray and a UV/optical focusing telescopes respectively which produce subarcsecond positions and multiwavelength light-curves for gamma-ray Burst (GRB) afterglows. BAT is a wide field-of-view (FOV) coded-aperture gamma-ray imager that produces arcminute GRB positions onboard within 10 seconds. The spacecraft executes a rapid autonomous slew that points the focusing telescopes at the BAT position in typically 50 seconds.

Synchrotron radiation: Type of nonthermal radiation caused by high-speed charged particles, such as electrons, emitting radiation as they are accelerated in a magnetic field. In accelerator physics, it is produced when electrons or positrons fly through deflecting magnets of ring accelerators or through wigglers or undulators. It is used for analyzing atomic and molecular structures in many natural sciences.

Time dilation: A prediction of the theory of relativity, closely related to the gravitational redshift. To an outside observer, a clock lowered into a strong gravitational field will appear to run slow.

Ultraluminous X-ray source (ULX): An ultraluminous X-ray source (ULX) is an astronomical source of X-rays that is not in the nucleus of a galaxy, and is more luminous than 10^{32} watt, brighter than the Eddington luminosity of a 10 solar-mass black hole. Typically there is about one ULX per galaxy in galaxies which host ULXs, but some galaxies contain many ULXs. The Milky Way does not contain an ULX. A survey of ULXs by Chandra observations shows that there is approximately one ULX per galaxy in galaxies which host ULXs. ULXs are found in all types of galaxies, including elliptical galaxies, but are more ubiquitous in star forming galaxies and in gravitationally interacting galaxies.

Visual binary: A binary star system in which both members are resolvable from Earth.

White dwarf: A star that has exhausted most or all of its nuclear fuel and has collapsed to a very small size (about the Earth's size). These stars are not heavy enough to generate the core temperatures required to fuse carbon in nucleosynthesis reactions. After it has become a red giant during its helium-burning phase, it will shed its outer layers to form a planetary nebula, leaving behind an inert core consisting mostly of carbon and oxygen. The white dwarf is supported only by electron degeneracy pressure. The maximum mass of a white dwarf, beyond which degeneracy pressure can no longer support it, is about 1.4 solar masses depending on its chemical composition.

XMM–Newton: The European Space Agency's large X-ray observatory, launched on December 10, 1999, which is capable of sensitive X-ray spectroscopic observations. XMM–Newton's name comes from the design of its mirrors, the highly nested X-ray Multi-Mirrors. XMM–Newton's highly eccentric orbit (with apogee of 114,000 km away from Earth and a perigee of 7000 km) has been chosen so that its instruments can work outside the radiation belts surrounding the Earth.

X-ray: Region of the electromagnetic spectrum corresponding to radiation of high frequency, corresponding to energies from 0.1 keV to 100 keV, and short wavelengths, far outside the visible spectrum.

X-ray binary: A binary star system in which a normal star is in orbit around a stellar remnant. The remnant accretes material from the normal star and produces X-rays in the process.

X-ray burster: X-ray source that radiates thousands of times more energy than our Sun, in short bursts that last only a few seconds. A neutron star in a binary system accretes matter onto its surface until temperatures reach the level needed for hydrogen fusion to occur. The result is a sudden period of rapid nuclear burning and release of energy.

References

1. M. Camenzind: *General Relativity – The Theory of Space and Time*, Lecture Notes (University Heidelberg 2004), see <http://www.lsw.uni-heidelberg.de/users/mcamenzi>
2. S.M. Carroll: *Spacetime and Geometry: An Introduction to General Relativity* (Addison-Wesley, New York 2003)
3. S. Chandrasekhar: *Hydrodynamic and Hydromagnetic Stability*, Int. Ser. of Monographs on Physics (Oxford Clarendon Press, Oxford 1961)
4. S. Chandrasekhar: *The Mathematical Theory of Black Holes* (Oxford University Press, Oxford 1983)
5. Norman K. Glendenning: *Compact Stars – Nuclear Physics, Particle Physics and General Relativity*, Astronomy and Astrophysics Library, 2nd edn (Springer-Verlag, Berlin, Heidelberg 2000)
6. W.G.H. Lewin, M. van der Klis: *Compact Stellar X-Ray Sources*, Cambridge Astrophysics Series **39** (Cambridge University Press, Cambridge 2006)
7. A.G. Lyne, F.G. Smith: *Pulsar Astronomy* (Cambridge University Press, Cambridge 1998)
8. R.N. Manchester, J.H. Taylor: *Pulsars* (Freeman, New York 1977)
9. F.C. Michel: *Theory of Neutron Star Magnetospheres* (Chicago University Press, Chicago 1991)
10. C. Misner, K. Thorne, J. Wheeler: *Gravitation* (Freeman, San Francisco 1973)
11. M. Nakahara: *Geometry, Topology and Physics*, 2nd edn (IOP, Bristol 2003)
12. C. Rovelli: *Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge 2004)
13. G.B. Rybicki, A.P. Lightman: *Radiation Processes* (Wiley Interscience, New York 1979)
14. B.F. Schutz: *A First Course in General Relativity* (Cambridge University Press, Cambridge 1986)
15. Stuart L. Shapiro, Saul A. Teukolsky: *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects* (John Wiley, New York 1983)
16. E.M. Sion, S. Vennes, H.L. Shipman: *White Dwarfs: Cosmological and Galactic Probes*, Astrophysics and Space Science Library **332** (Springer-Verlag, Berlin 2005)
17. H. Stephani: *Relativity – An Introduction to Special and General Relativity*, 3rd edn (Cambridge University Press, Cambridge 2004)
18. N. Straumann: *General Relativity and Relativistic Astrophysics*, 2nd edn (Springer-Verlag, Berlin 2000)
19. F. Weber: *Pulsars as Astrophysical Laboratories for Nuclear and Particle Physics* (IOP, Bristol 1999)
20. S. Weinberg: *Gravitation and Cosmology* (Wiley, New York 1972)

21. M.A. Abramowicz, B. Czerny, J.-P. Lasota, E. Szuszkiewicz: *Slim accretion disks*, ApJ **332**, 646 (1988)
22. M.A. Abramowicz, X. Chen, S. Kato, J.-P. Lasota, O. Regev: *Thermal equilibria of accretion disks*, ApJ **438**, L37 (1995)
23. M.A. Abramowicz, W. Kluzniak, J.P. Lasota: *No observational proof of the black-hole event-horizon*, A&A **396**, L31 (2002); arXiv:astro-ph/0207270
24. F.A. Aharonian, M.A. Atoyan: *Nonthermal radiation of the Crab Nebula*, in Proc. *Neutron Stars and Pulsars: Thirty Years after the Discovery*, ed. N. Shibasaki et al. (Universal Academy Press, Tokoyo 1998), p 439; arXiv:astro-ph/9803091
25. F.A. Aharonian et al.: *The Crab Nebula and pulsar between 500 GeV and 80 TeV: Observations with the HEGRA stereoscopic air Cherenkov*, ApJ **614**, 897 (2004); astro-ph/0407118
26. A. Akmal, V.R. Pandharipande, D.G. Ravenhall: *Equation of state of nucleon matter and neutron star structure*, Phys. Rev. C **58**, 1804 (1998)
27. C. Alcock, E. Farhi, A. Olinto: *Strange stars*, ApJ **310**, 261 (1986)
28. M. Alcubierre, B. Brügmann: *Simple excision of a black hole in 3+1 numerical relativity*, Phys. Rev. D **62**, 14011 (2001)
29. M. Alford: *Color superconducting quark matter*, Ann. Rev. Nucl. Part. Sci. **51**, 131 (2001)
30. M. Alford: *Dense quark matter in nature*, nucl-th/0312007 (2003)
31. M. Alford, S. Reddy: *Compact stars with color superconducting quark matter*, Phys. Rev. D **67**, 74024 (2004)
32. M. Alford, M. Braby, M. Paris, S. Reddy: *Hybrid stars that masquerade as neutron stars*, arXiv:nucl-th/0411016 (2005)
33. A.M. Anile: *Relativistic Fluids and Magnetofluids*, Cambridge University Press (Cambridge 1989)
34. M. Ansorg, A. Kleinwächter, R. Meinel: *Highly accurate calculation of rotating neutron stars*, A&A **381**, L49 (2002)
35. M. Ansorg, A. Kleinwächter, R. Meinel: *Highly accurate calculation of rotating neutron stars: Detailed description of the numerical methods*, A&A **405**, 711 (2003); arXiv:astro-ph/0103173
36. L. Anton, O. Zanotti, J.A. Miralles et al.: *Numerical 3+1 general relativistic magnetohydrodynamics: A local characteristics approach*, ApJ **637**, 296 (2005); arXiv:astro-ph/0506063
37. S. Appl, M. Camenzind: *The structure of relativistic MHD jets: a solution to the nonlinear Grad-Shafranov equation*, A&A **274**, 699 (1993)
38. P.J. Armitage: *Turbulence and angular momentum transport in global accretion disk simulation*, ApJ **501**, L189 (1998)
39. P.J. Armitage, C.S. Reynolds: *The variability of accretion on to Schwarzschild black holes from turbulent magnetized discs*, MNRAS **339**, 1041 (2003)
40. R. Arnowitt, S. Deser, C.W. Misner: *The dynamics of general relativity*, in *Gravitation: An Introduction to Current Research*, ed. L. Witten (John Wiley, New York 1962), p 227
41. J. Arons: *Photon bubbles – Overstability in a magnetized atmosphere*, ApJ **388**, 561 (1992)
42. P. Arras, O. Blaes, N.J. Turner: *Quasi-periodic oscillations from magnetorotational turbulence*, arXiv:astro-ph/0602275
43. Yu.V. Artemova, G.S. Bisnovatyi-Kogan, I.V. Igumenshchev, I.D. Novikov: *Accretion disks with optical depth transition and advection*, Mem. Soc. Astron. Italiana **76**, 84 (2005)

44. A. Ashtekar: *New variables for classical and quantum gravity*, Phys. Rev. Lett. **57**, 2244 (1986)
45. A. Ashtekar: *New Hamiltonian formulation of general relativity*, Phys. Rev. D **36**, 1586 (1987)
46. L. Baiotti, I. Hawke, P. Montero, F. Löffler, L. Rezzolla, N. Stergioulas, T. Font, Ed. Seidel: *Three-dimensional relativistic simulations of rotating neutron star collapse to a Kerr black hole*, Phys. Rev. D **71**, 024035 (2005)
47. S. Balberg, I. Lichtenstadt, G.B. Cook: *Roles of hyperons in neutron stars*, ApJSupp **121**, 515 (1999)
48. S. Balbus, J.F. Hawley: *A powerful local shear instability in weakly magnetized disks. I – Linear analysis. II – Nonlinear evolution*, ApJ **376**, 214 (1991)
49. S. Balbus, J.F. Hawley: *Instability, turbulence, and enhanced transport in accretion disks*, Rev. Mod. Phys. **70**, 1 (1998)
50. S. Balbus: *Enhanced angular momentum transport in accretion disks*, ARA&A **41**, 555 (2003)
51. S. Balbus: *Turbulent energy transport in nonradiative accretion flows*, ApJ **600**, 865 (2004)
52. M. Baldo, I. Bombaci, G.F. Burgio: *Microscopic nuclear equation of state with three-body forces and neutron star structure*, A&A **328**, 274 (1997)
53. D. Balsara: *Total variation diminishing scheme for relativistic magnetohydrodynamics*, ApJSupp **132**, 83 (2001)
54. F. Banyuls, J.A. Font, J.M. Ibanez, J.M. Martí, J.A. Miralles: *Numerical 3+1 general relativistic hydrodynamics: A local characteristic approach*, ApJ **476**, 221 (1997)
55. J.M. Bardeen: *Stability of circular orbits in stationary, axisymmetric space-times*, ApJ **161**, 103 (1970)
56. J.M. Bardeen: *A variational principle for rotating stars in general relativity*, ApJ **162**, 171 (1970)
57. J.M. Bardeen, R.V. Wagoner: *Relativistic disks. I. Uniform rotation*, ApJ **167**, 359 (1971)
58. J. M. Bardeen, B. Carter, S.W. Hawking: *The four laws of black hole mechanics*, Comm. Math. Phys. **31**, 161 (1973)
59. M.A. Barstow, H.E. Bond, J.B. Holberg et al.: *Hubble Space Telescope spectroscopy of the Balmer lines in Sirius B*, MNRAS **362**, 1134 (2005); astro-ph/0506600
60. A.J. Barth, J.E. Greene, L.C. Ho: *Dwarf Seyfert 1 nuclei and the low-mass end of the $M_{BH} - \sigma$ relation*, ApJ **619**, L151 (2005)
61. T.w. Baumgarte, S.L. Shapiro: *Evolution of rotating supermassive stars to the onset of collapse*, ApJ **526**, 941 (1999)
62. T.W. Baumgarte, S.L. Shapiro: *General relativistic magnetohydrodynamics for the numerical construction of dynamical spacetimes*, ApJ **585**, 921 (2003)
63. A. Bauswein: *Struktur schnell rotierender Neutronensterne*, Diploma thesis, Technical University Darmstadt 2006
64. G.C. Baym, D. Pethick, P. Sutherland: *The ground state of matter at high densities: Equation of state and stellar models*, ApJ **170**, 299 (1971)
65. G.C. Baym, H. Bethe, C.J. Pethick: *Neutron star matter*, Nucl. Phys. A **175**, 225 (1979)
66. W. Becker, G.G. Pavlov: *The Milky Way – pulsars and isolated neutron stars*, in *The Century of Space Science*, eds. J. Bleeker, J. Geiss, M. Huber (Kluwer Academic Publishers 2002); arXiv:astro-ph/0208356
67. I. Bednarek, R. Manka: *The influence of the strength of hyperon–hyperon interactions on neutron star properties*, arXiv:hep-ph/0506059 (2005)

68. T. Belloni: *Black hole states: Accretion and jet ejection*, astro-ph/0504185 (2005)
69. T. Belloni, M. Mendez, J. Homan: *The distribution of kHz QPO frequencies in bright LMXBs*, A&A **437**, 209 (2005)
70. E. Berti, N. Stergioulas: *Approximate matching of analytic and numerical solutions for rapidly rotating neutron stars*, MNRAS **350**, 1416 (2004)
71. E. Berti, F. Frances, A. Maniopoulou, M. Bruni: *Rotating neutron stars: An invariant comparison of approximate and numerical space-time models*, MNRAS **358**, 923 (2005)
72. V.S. Beskin, I.V. Kuznetsova: *On the Blandford-Znajek mechanism of the energy loss of a rotating black hole*, Il Nuovo Cimento B **115**, 795 (2000)
73. G.S. Bisnovatyi-Kogan, S.I. Blinnikov: *Disk accretion onto a black hole at subcritical luminosity*, A&A **59**, 111 (1977)
74. O. Blaes: *Physics fundamentals of luminous accretion disks around black holes*, in *Accretion discs, jets and high energy phenomena in astrophysics*, ed. V. Beskin, G. Henri, F. Menard et al, Les Houches Summer School **78**, 137 (2004)
75. O.M. Blaes, A. Socrates: *Local dynamical instabilities in magnetized, radiation pressure-supported accretion disks*, ApJ **553**, 987 (2001)
76. O.M. Blaes, A. Socrates: *Local radiative hydrodynamic and magnetohydrodynamic instabilities in optically thick media*, ApJ **596**, 509 (2003)
77. L. Blanchet, T. Damour, B.L. Iyer, C.M. Will, A.G. Wiseman: *Gravitational-radiation damping of compact binary systems to second post-Newtonian order*, Phys. Rev. Lett. **74**, 3515 (1995)
78. L. Blanchet: *Gravitational radiation from post-Newtonian sources and inspiralling compact binaries*, Living Reviews in Relativity **Irr-2002-3** (2002)
79. R.D. Blandford, S.A. Teukolsky: *Arrival-time analysis for a pulsar in a binary system*, ApJ **205**, 580 (1976)
80. R.D. Blandford, R.L. Znajek: *Electromagnetic extraction of energy from Kerr black holes*, MNRAS **179**, 433 (1977)
81. D. Blaschke, D. Voskresensky, H. Grogorian: *Cooling of neutron stars with color superconducting quark cores*, arXiv:hep-ph/0510368 (2005)
82. I. Bombaci, M. Prakash, Ainsworth, J.M. Lattimer: *Physics of neutron star interiors*, Phys. Rep. **280**, 1 (1997)
83. S. Bonazzola, E. Gourgoulhon, M. Salgado, J.A. Marck: *Axisymmetric rotating relativistic bodies: A new numerical approach for exact solutions*, A&A **278**, 421 (1993)
84. S. Bonazzola, E. Gourgoulhon: *A virial identity applied to relativistic stellar models*, Class. Quantum Grav. **11**, 1775 (1994)
85. S. Bonazzola, E. Gourgoulhon, J.A. Marck: *Numerical approach for high precision 3-D relativistic star models*, Phys. Rev. D **58**, 104020 (1998); arXiv:astro-ph/9803086
86. R.H. Boyer, R.W. Lindquist: *Maximal analytic extension of the Kerr metric*, J. Math. Phys. **8**, 265 (1967)
87. S. Brandt, B. Brügmann: *A simple construction of initial data for multiple black holes*, Phys. Rev. Lett. **78**, 3606 (1997)
88. S. Brinkmann: *MHD-instabilities in accretion disks*, Diploma thesis, University Heidelberg 2004
89. S. Brinkmann, M. Camenzind, J. Gracia: *On the global structure of magnetohydrodynamical unstable non-radiative accretion discs*, submitted for publication (2006)
90. M. Brio, C.C. Wu: *An upwind differencing scheme for the equations of ideal magnetohydrodynamics*, J. Comp. Phys. **75**, 400 (1988)
91. G.E. Brown: *High-density equation of state*, Nature **336**, 519 (1988)

92. M. Buballa: *NJL-model analysis of dense quark matter*, Phys. Rep. **407**, 205 (2005)
93. V. Burwitz, F. Haberl, R. Neuhauser et al.: *The thermal radiation of the isolated neutron star RX J1856.5–3754 observed with Chandra and XMM–Newton*, A&A **399**, 1109 (2003)
94. E.M. Butterworth, and J.R. Ipser: *On the structure and stability of rapidly rotating fluid bodies in general relativity. I – The numerical method for computing structure and its application to uniformly rotating homogeneous bodies*, ApJ. **204**, 200 (1976)
95. A. Cadez, C. Fanton, M. Calvani: *Line emission from accretion discs around black holes: The analytic approach*, New Astron. **3**, 647 (1998)
96. M. Camenzind: *Hydromagnetic flows from rapidly rotating compact objects. I – Cold relativistic flows from rapid rotators*, A&A **162**, 32 (1986)
97. M. Camenzind: *Hydromagnetic flows from rapidly rotating compact objects. II – The relativistic axisymmetric jet equilibrium*, A&A **184**, 341 (1987)
98. M. Camenzind: *Magnetohydrodynamics of rotating black holes*, in: *Relativistic Astrophysics*, ed. H. Riffert, H. Ruder, H.-P. Nollert, F.W. Hehl, p 82 (Vieweg, Braunschweig 1998)
99. M. Camenzind, R. Khanna: *Magnetohydrodynamic processes near rapidly rotating compact objects*, Il Nuovo Cimento B **115**, 815 (2000)
100. M. Camenzind: *The black hole environments*, in *Accretion discs, jets and high energy phenomena in astrophysics*, ed. V. Beskin, G. Henri, F. Menard et al., Les Houches Summer School **78**, 405 (2004)
101. M. Camenzind: *Relativistic outflows from active galactic nuclei*, Mem. Soc. Astron. Italiana **76**, 98 (2005)
102. M. Camenzind: *Cosmic black holes – From stellar to supramassive black holes in galaxies*, Ann. Physik **15**, 60 (2006)
103. M. Camenzind, A. Boucher: *Les noyaux actifs de galaxies*, Lecture Notes in Physics **46**, (Springer-Verlag, Berlin 1997)
104. B.J. Carr: *Primordial black holes – Recent developments*, arXiv:astro-ph/0504034 (2005)
105. B. Carter: *Hamilton–Jacobi and Schrödinger separable solutions of Einstein’s equations*, Comm. Math. Phys. **10**, 280 (1968)
106. B. Carter: *The commutation property of a stationary, axisymmetric system*, Comm. Math. Phys. **17**, 233 (1970)
107. B. Carter: *The general theory of mechanical, electromagnetic and thermodynamic properties of black holes*, in: *General Relativity: An Einstein Centenary Survey*, ed. S.W. Hawking, W. Israel, p 294 (Cambridge Univ. Press, Cambridge 1979)
108. J. Casares: *The mass-spectrum of X-Ray binaries*, arXiv:astro-ph/0503071 (2005)
109. C. Catteon, T. Faber, Matt Visser: *Gravastars must have anisotropic pressure*, arXiv:gr-qc/0505137 (2005)
110. P. Cerdá-Durán, G. Faye, H. Dimmelmeier, J.A. Font, J.M. Ibanez, E. Mueller, G. Schaeffer: *CFC+: Improved dynamics and gravitational waveforms from relativistic core collapse simulations*, A&A **439**, 1033 (2005); arXiv:astro-ph/041261
111. G. Chabrier, P. Brassard, G. Fontaine, D. Saumon: *Cooling sequences and color-magnitude diagrams for cool white dwarfs with hydrogen atmospheres*, ApJ **543**, 216 (2000); arXiv:astro-ph/0006363
112. D. J. Champion, D. R. Lorimer, M. A. McLaughlin, J. M. Cordes, Z. Arzoumanian, J. M. Weisberg, J. H. Taylor: *PSR J1829+2456: A relativistic binary pulsar*, MNRAS **350**, L61 (2004)
113. S. Chandrasekhar: *The maximum mass of ideal white dwarfs*, ApJ **74**, 81 (1931)

114. S. Chandrasekhar: *The equilibrium of distorted polytropes. I. The rotational problem*, MNRAS **93**, 390 (1933)
115. S. Chandrasekhar: *The dynamical instability of gaseous masses approaching the Schwarzschild limit in general relativity*, ApJ **140**, 417 (1964)
116. G. Chapline: *Dark energy stars*, arXiv:astro-ph/0503200 (2005)
117. G. Chapline, E. Hohlfeld, B. Laughlin, D. I. Santiago: *Quantum phase transitions and the breakdown of classical general relativity*, Int. J. Mod. Phys. A **18**, 3587 (2003); arXiv:gr-qc/0012094
118. G. Chapline, E. Hohlfeld, R. B. Laughlin and D. I. Santiago, Quantum phase transitions and the breakdown of classical general relativity, Int. J. Mod. Phys. A 18 3587 (2003) [arXiv:gr-qc/0012094].
119. A. Chodos et al.: *New extended model of hadrons*, Phys. Rev. D **9**, 3471; *Baryon structure in the bag theory*, Phys. Rev. D **10**, 2599 (1974)
120. <http://seesar.lbl.gov/anag/chombo/>
121. R. Coker, F. Melia: *Hydrodynamical accretion onto Sagittarius A* from distributed point sources*, ApJ **488**, L149 (1997)
122. J. Contopoulos, D. Kazanas, C. Fendt: *The axisymmetric pulsar magnetosphere*, ApJ **511**, 351 (1999)
123. G.B. Cook, S.L. Shapiro, S.A. Teukolsky: *Spin-up of a rapidly rotating star by angular momentum loss – Effects of general relativity*, ApJ **398**, 203 (1992)
124. G.B. Cook, S.L. Shapiro, S.A. Teukolsky: *Rapidly rotating neutron stars in general relativity. Realistic equations of state*, ApJ **424**, 828 (1994)
125. S. Corbel: *Large scale jets in microquasars*, Memorie della Societa Astronomica It. **76**, 73 (2005)
126. J. Cottam, F. Paerels, M. Mendez: *Gravitationally redshifted absorption lines in the X-ray burst spectra of a neutron star*, Nature **420**, 51 (2003)
127. C.T. Cunningham: *The effects of redshifts and focusing on the spectrum of an accretion disk around a Kerr black hole*, Ap. J. **202**, 788 (1975)
128. T. Damour, N. Deruelle: Ann. Inst. H. Poincare A **44**, 263 (1986)
129. T. Damour, J.H. Taylor: *Strong-field tests of relativistic gravity and binary pulsars*, Phys. Rev. D **45**, 1840 (1992)
130. S.W. Davis, O:M. Blaes, I. Hubeny, N.J. Turner: *Relativistic accretion disk models of high state black hole X-ray binary spectra*, ApJ **621**, 372 (2005)
131. A. Dedner, F. Kemm, D. Kröner et al.: *Hyperbolic divergence cleaning for the MHD equations*, J. Comp. Phys. **175**, 645 (2002)
132. L. Del Zanna, N. Bucciantini, P. Londrillo: *An efficient shock-capturing central-type scheme for multidimensional relativistic flows. II. Magnetohydrodynamics*, A&A **400**, 397 (2003)
133. J.-P. De Villiers, J.F. Hawley: *A numerical method for general relativistic magnetohydrodynamics*, ApJ **589**, 458 (2003)
134. J.-P. De Villiers, J.F. Hawley, J.H. Krolik: *Magnetically driven accretion flows in the kerr metric. I. Models and overall structure*, ApJ **599**, 1238 (2003)
135. J.-P. De Villiers, J.F. Hawley, J.H. Krolik, S. Hirose: *Magnetically driven accretion in the Kerr metric. III. Unbound outflows*, ApJ **620**, 878 (2005)
136. F. Douchin, P. Haensel: *A unified equation of state of dense matter and neutron star structure*, A&A **380**, 151 (2001)
137. H.P. Duerr: *Relativistic effects in nuclear forces*, Phys. Rev. **103**, 469 (1956)
138. M.D. Duez, Y.T. Liu, S.L. Shapiro, B.C. Stephens: *Relativistic magnetohydrodynamics in dynamical spacetimes: numerical methods and tests*, Phys. Rev. D **72**, 024028 (2005); astro-ph/0503420

139. R. Durrer, N. Straumann: *Some Applications of the 3+1 Formalism of General Relativity*, Helv. Phys. Acta **61**, 1027 (1988)
140. I. Dymnikova: *Spherically symmetric space-time with regular de Sitter center*, Int. J. Mod. Phys. D **12**, 1015 (2003)
141. A. Einstein: *Die Grundlage der allgemeinen Relativitätstheorie*, Ann. Phys. **49**, 769 (1916)
142. A. Einstein, Sitzber. Preuss. Akad. Wiss. Berlin 688 (1916) Näherungsweise Integration der Feldgleichungen der Gravitation
143. A. Einstein, Sitzber. Preuss. Akad. Wiss. Berlin 154 (1918) Über Gravitationswellen
144. R. Epstein: *The binary pulsar: Post-Newtonian timing effects*, ApJ **216**, 92 (1977)
145. F.J. Ernst: *A new formulation of the axially symmetric gravitational field problem*, Phys. Rev. **167**, 1175 (1968)
146. F.J. Ernst: *A new formulation of the axially symmetric gravitational field problem II*, Phys. Rev. **168**, 1415 (1968)
147. A.A. Esin, J.E. McClintock, R. Narayan: *Advection-dominated accretion and the spectral states of black hole X-ray binaries: Application to Nova MUSCAE 1991*, ApJ **489**, 865 (1997)
148. A.A. Esin, J.E. McClintock et al.: *Modeling the low state spectrum of the X-ray Nova XTE J1118+480*, ApJ **555**, 483 (2001)
149. C. Evans, J.F. Hawley: *Simulation of magnetohydrodynamic flows – A constrained transport method*, ApJ **332**, 659 (1988)
150. A.C. Fabian, S. Vaughan, K. Nandra et al.: *A long hard look at MCG–6–30–15 with XMM–Newton*, MNRAS **335**, L1 (2002)
151. H. Falcke, F. Melia, E. Agol: *Viewing the shadow of the black hole at the Galactic center*, ApJ **528**, L13 (2000)
152. S.A.E.G. Falle: *Rarefaction shocks, shock errors, and low order of accuracy in ZEUS*, ApJ **577**, L123 (2002)
153. C. Fanton, M. Calvani, F. de Felice, A. Cadez: *Detecting accretion disks in active galactic nuclei*, PASJ **49**, 159 (1997)
154. A.J. Faulkner, M.Kramer, A.G.Lyne et al.: *PSR J1756–2251: a new relativistic double neutron star system*, ApJ **618**, L119 (2004)
155. R. Bender, T.M. Belloni, E. Gallo: *Towards a unified model for black hole X-ray binary jets*, MNRAS **335**, 1105 (2004)
156. C. Fendt: *Stationary models of relativistic magnetohydrodynamic jets*, Proc. 3rd Int. Sakharov Conf. on Physics, **2**, 315 (Scientific World 2003)
157. C. Fendt, F. Memola: *Collimating, relativistic, magnetic jets from rotating disks. The axisymmetric field structure of relativistic jets and the example of the M87 jet*, A&A **365**, 631 (2001)
158. C. Fendt, F. Memola: *Stationary relativistic magnetic jets from black holes*, ApSS **276**, 297 (2001)
159. J.A. Font: *Numerical hydrodynamics in general relativity*, Living Reviews in Relativity **Irr-2003-4** (2003)
160. R.H. Fowler: *Dense matter*, MNRAS **87**, 114 (1926)
161. J.L. Friedman, L. Parker, J.R. Ipser: *Rapidly rotating neutron star models*, ApJ **304**, 115 (1986)
162. B.L. Friman, O.V. Maxwell: *Neutrino emissivities of neutron stars*, ApJ **232**, 541 (1979)
163. B. Fryxell, K. Olson, P. Ricker et al.: *FLASH: An adaptive mesh hydrodynamics code for modeling astrophysical thermonuclear flashes*, ApJS **131**, 273 (2000)
164. C.F. Gammie: *Photon bubbles in accretion discs*, MNRAS **297**, 929 (1998)

165. C.F. Gammie: *The magnetorotational instability in the Kerr metric*, ApJ **614**, 309 (2004)
166. C.F. Gammie, J.C. McKinney, G. Toth: *HARM: A numerical scheme for general relativistic magnetohydrodynamics*, ApJ **589**, 444 (2003)
167. K. Gebhardt, R. Bender, G. Bower et al.: *A relationship between nuclear black hole mass and velocity dispersion*, ApJ **539**, L13 (2000)
168. D.M. Gelino et al.: *The inclination angle and mass of the black hole in XTE J1118+480*, astro-ph/0601409
169. R. Genzel, R. Schödel, T. Ott et al.: *Near-infrared flares from accreting gas around the supermassive black hole at the Galactic Centre*, Nature **425**, 934 (2003)
170. U.H. Gerlach: *Equation of state at supranuclear densities and the existence of a third family of superdense stars*, Phys. Rev. **172**, 1325 (1968)
171. A.M. Ghez, S.A. Wright, K. Matthews et al.: *Variable infrared emission from the supermassive black hole at the center of the Milky Way*, ApJ **601**, L159 (2003)
172. B. Giacomazzo, L. Rezzolla: *The exact solution of the Riemann problem in relativistic MHD*, J. Fluid Mech. **562**, 223 (2006); gr-qc/0507102
173. A.M. Ghez, S. Salim, S.D. Hornstein et al.: *Stellar orbits around the Galactic Center black hole*, ApJ **620**, 744 (2005)
174. N.K. Glendenning, N.K. 1985, *Neutron stars are giant hypernuclei?*, ApJ **293**, 470 (1985)
175. N.K. Glendenning: *Vacuum polarization effects on nuclear matter and neutron stars*, Nucl. Phys. A **493**, 521 (1989)
176. P. Goldreich, W.H. Julian: *Pulsar electrodynamics*, ApJ **157**, 869 (1969)
177. E. Gourgoulhon: *A 3+1 perspective on null hypersurfaces and isolated horizons*, arXiv:gr-qc/0503113 (2005)
178. J.E. Goldston, E. Quataert, I.V. Igumenshchev: *Synchrotron radiation from radiatively inefficient accretion flow simulations: applications to Sagittarius A**, ApJ **621**, 785 (2005)
179. D. Gondek-Rosinska, E. Gourgoulhon: *Jacobi-like bar mode instability of relativistic rotating bodies*, Phys. Rev. D **66** 044021 (2002)
180. E. Gourgoulhon, S. Bonazzola: *A formulation of the virial theorem in general relativity*, Class. Quantum Grav. **11**, 443 (1994)
181. E. Gourgoulhon, P. Haensel, R. Livine, E. Paluch, S. Bonazzola, J.-A. Marck: *Fast rotation of strange stars*, A&A **349**, 851 (1999)
182. J. Gracia, J. Peitz, Ch. Keller, M. Camenzind, M.: *Evolution of bimodal accretion flows*, MNRAS **344**, 468 (2003)
183. L.J. Greenhill, D.R. Jiang, J.M. Moran et al.: *Detection of a subparsec diameter disk in the nucleus of NGC 4258*, ApJ **440**, 619 (1995)
184. A. Gurevich, V. Beskin, Y. Istomin: *Physics of Pulsar Magnetosphere* (Cambridge University Press, Cambridge 1993)
185. M.E. Gusakov, A.D. Kaminker, D.G. Yakovlev, O.Y. Gnedin: *Cooling of Akmal-Pandharipande-Ravenhall neutron star models*, arXiv:astro-ph/0507560 (2005)
186. P. Haensel, J.L. Zdunik, R. Schaeffer: *Strange quark stars*, A&A **160**, 121 (1986)
187. P. Haensel: *Equation of state of dense matter and maximum mass of neutron stars*, in *Final Stages of Stellar Evolution*, ed. C. Motch, J.-M. Hameury, EAS Publ. Series (EDP Sciences), pp 249-284 (2003)
188. P. Haensel, A.Y. Potekhin: *Analytical representations of unified equations of state of neutron-star matter*, A&A **428**, 191 (2004)
189. N. Häring, H.-W. Rix: *On the black hole mass–bulge mass relation*, ApJ **604**, L89 (2004)

190. T. Hamada, E.E. Salpeter: *Models for zero-temperature stars*, ApJ **134**, 683 (1961)
191. B.M.S. Hansen, J. Liebert: *Cool white dwarfs*, ARA&A **41**, 465 (2003)
192. B.M.S. Hansen et al.: *HST observations of WD cooling sequence in M4*, ApJSup **155**, 551 (2004); arXiv:astro-ph/0401443
193. H.C. Harris, J.A. Munn, M. Kilic et al.: *The white dwarf luminosity function from SDSS imaging data*, arXiv:astro-ph/0510820
194. A. Harten, P.D. Lax, B. van Leer: *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Review **25**, 35 (1983)
195. J.B. Hartle, Kip S. Thorne: *Slowly rotating relativistic stars. II. Models for neutron stars and supermassive stars*, ApJ **153**, 807 (1968)
196. M. Haugan: *Post-Newtonian arrival-time analysis for a pulsar in a binary system*, Ap. J. **296**, 1 (1985)
197. S. Hawking: *Black hole explosions?* Nature **248**, 30 (1974)
198. S. Hawking: *Black holes and thermodynamics*, Phys. Rev. D **13**, 191 (1976)
199. J.F. Hawley: *Global magnetohydrodynamical simulations of accretion tori*, ApJ **528**, 462 (2000)
200. J.F. Hawley, J.-P. De Villiers: *General relativistic magnetohydrodynamic simulations of black hole accretion disks*, Progr. Theoret. Phys. Supp **155**, 132 (2004)
201. J.R. Herrnstein, L.J. Greenhill, J.M. Moran: *VLBA continuum observations of NGC 4258: Constraints on an advection-dominated accretion flow*, ApJ **497**, L69 (1998)
202. S. Hirose, J.H. Krolik, J.P. De Villiers, J.F. Hawley: *Magnetically driven accretion flows in the Kerr metric. II. Structure of the magnetic field*, ApJ **606**, 1083 (2004)
203. G. t'Hooft: *Dimensional reduction in quantum gravity*, gr-qc/9310026 (1993)
204. I. Hubeny, E. Agol, O. Blaes, J. Krolik: *Non-LTE models and theoretical spectra of accretion disks in active galactic nuclei. III. Integrated spectra for hydrogen–helium disks*, ApJ **533**, 710 (1999)
205. A. Hujeirat, R. Rannacher: *On the efficiency and robustness of implicit methods in computational astrophysics*, NewAR **45**, 425 (2001)
206. A. Hujeirat: *A model for electromagnetic extraction of rotational energy and formation of accretion-powered jets in radio galaxies*, A&A **416**, 423 (2004)
207. A. Hujeirat: *A method for enhancing the stability and robustness of explicit schemes in CFD*, NewA **10**, 173 (2005)
208. A. Hujeirat, M. Camenzind: *Truncated disks – advective tori solutions around BHs. I. The effects of conduction and enhanced Coulomb coupling*, A&A **362**, L41 (2000)
209. A. Hujeirat, M. Camenzind, A. Burkert: *Comptonization and synchrotron emission in 2D accretion flows. I. A new numerical solver for the Kompaneets equation*, A&A **386**, 757 (2002)
210. R. Ibata et al.: *Faint, Moving objects in the Hubble Deep Field: Components of the dark halo?*, ApJ **524**, L95 (1999)
211. S. Ichimaru: *Bimodal behavior of accretion disks: Theory and application to Cygnus X-1 transitions*, ApJ **214**, 840 (1977)
212. I.V. Igumenshchev, R. Narayan, M.A. Abramowicz: *Three-dimensional magnetohydrodynamic simulations of radiatively inefficient accretion flows*, ApJ **592**, 1042 (2003)
213. J.R. Ipser, L. Lindblom: *On the adiabatic pulsations of accretion disks and rotating stars*, ApJ **379**, 285 (1991)
214. T. Islam, S. Balbus: *The dynamics of the magnetoviscous instability*, astro-ph/0504666
215. N. Iwamoto: *Quark beta decay and the cooling of neutron stars*, Phys. Rev. Lett. **44**, 1637 (1980)
216. B.A. Jacoby, A. Hotan, M. Bailes et al.: *The mass of a millisecond pulsar*, arXiv:astro-ph/0507420 (2005)

217. H.-T. Janka: *Neutron star formation and birth properties*, arXiv:astro-ph/0404200
218. M.H. Johnson, E. Teller: *Classical field theory of nuclear forces*, Phys. Rev. **98**, 783 (1955)
219. S. Kato, J. Fukue, S. Mineshige: *Black-hole accretion disks*, in: *Black-Hole Accretion Disks* (Kyoto University Press, Kyoto 1998)
220. R.P. Kerr: *Gravitational field of a spinning mass as an example of algebraically special metrics*, Phys. Rev. Lett. **11**, 237 (1963)
221. R.P. Kerr, A. Schild: *A new Class of Vacuum Solutions of the Einstein Field Equations*, IV Centenario Della Nascita di Galileo Galilei, 1564-1964. Pubblicazioni del Comitato Nazionale per le Manifestazioni Celebrazive. Ed. G. Barbera (Firenze), p 222 (1965)
222. R. Khanna, M. Camenzind: *The ω dynamo in accretion disks of rotating black holes*, A&A **307**, 665 (1996); A&A **313**, 108 (1996)
223. R. Khanna: *Der gravitomagnetische Dynamoeffekt in Akkretionsscheiben Schwarzer Löcher*, PhD Thesis, University of Heidelberg 1993
224. R. Khanna: *On the magnetohydrodynamic description of a two-component plasma in the Kerr metric*, MNRAS **294**, 673 (1998)
225. R. Khanna: *Generation of magnetic fields by a gravitomagnetic plasma battery*, MNRAS **295**, L6 (1998)
226. M. Kilic, Ted von Hippel, F. Mullally et al.: *The Mystery deepens: Spitzer observations of cool white dwarfs*, arXiv:astro-ph/0601305
227. S.J. Kleinman, H.C. Harris, D.J. Eisenstein et al.: *A catalog of spectroscopically identified white dwarf stars in the first data release of the Sloan Digital Sky Survey*, ApJ **607**, 426 (2004)
228. H. Komatsu, Y. Eriguchi, I. Hachisu: *Rapidly rotating general relativistic stars. I – Numerical method and its application to uniformly rotating polytropes*, MNRAS **237**, 355 (1989); *Rapidly rotating general relativistic stars. II – Differentially rotating polytropes*, MNRAS **239**, 153 (1989)
229. S.J. Kleinman et al. 2004, *A catalog of spectroscopically identified white dwarf stars from the first data release of the SDSS*, ApJ **607**, 426 (2004); arXiv:astro-ph/0402209
230. S. Koide, K. Shibata, T. Kudoh: *GRMHD simulations of jets from black hole accretion disks*, ApJ **495**, L63 (1998)
231. S. Koide, K. Shibata, T. Kudoh: *Relativistic jet formation from black hole magnetized accretion disks: Method, tests, and applications of a general relativistic magnetohydrodynamic numerical code*, ApJ **522**, 727 (1999)
232. S. Koide, K. Shibata, T. Kudoh, D.L. Meier: *Extraction of black hole magnetic energy by a magnetic field and the formation of relativistic jets*, Science **295**, 1688 (2002)
233. A.V. Koldoba, O.A. Kusnetsov, G. Ustyugova: *An approximate Riemann solver for relativistic MHD*, MNRAS **333**, 932 (2002)
234. A. Komar: *Covariant conservation laws in general relativity*, Phys. Rev. **113**, 934 (1959)
235. H. Komatsu, Y. Eriguchi, I. Hachisu: *Rapidly rotating general relativistic stars. II – Differentially rotating polytropes*, MNRAS **239**, 153 (1989)
236. S. Komissarov: *A Godunov-type scheme for relativistic magnetohydrodynamics*, MNRAS **303**, 343 (1999)
237. S. Komissarov: *On the properties of time-dependent, force-free, degenerate electrodynamics*, MNRAS **336**, 759 (2002)
238. S. Komissarov: *Electrodynamics of black hole magnetospheres*, MNRAS **350**, 407 (2004)
239. S. Komissarov: *General relativistic MHD simulations of monopole magnetospheres of black holes*, MNRAS **350**, 1431 (2004)

240. S. Komissarov: *Observations of the Blandford–Znajek and the MHD Penrose processes in computer simulations of black hole magnetospheres*, MNRAS **359**, 801 (2004); arXiv:astro-ph/0501599
241. S. Komissarov: *Simulation of axisymmetric magnetospheres of neutron stars*, MNRAS **367**, 19 (2006)
242. D. Koester: *White dwarfs: Recent developments*, A&ARev **11**, 33 (2002)
243. D. Koester, G. Chanmugam: *Physics of white dwarf stars*, Progr. Theor. Phys. **53**, 837 (1990)
244. M. Kramer, D.R. Lorimer, A.G. Lyne et al.: *Testing GR with the double pulsar: Recent results*, arXiv:astro-ph/0503386
245. M. Kramer, A.G. Lyne, M. Burgay et al.: *The double pulsar – A new testbed for relativistic gravity*, arXiv:astro-ph/0405179
246. J.H. Krolik, J.F. Hawley, S. Hirose: *Magnetically driven accretion flows in the Kerr metric. IV. Dynamical properties of the inner disk*, ApJ **622**, 1008 (2005)
247. L.D. Landau: *On the theory of stars*, Phys. Zeitschr. Sowjetunion **1**, 285 (1932)
248. J.M. Lattimer, D.F. Swesty: *An effective equation of state for hot dense matter*, Nucl. Phys. A **535**, 331 (1991)
249. J.M. Lattimer, M. Prakash: *Nuclear matter and its role in supernovae, neutron stars and compact object binary mergers*, Phys. Reports **333**, 121 (2000)
250. J.M. Lattimer, M. Prakash: *The physics of neutron stars*, Science **304**, 536 (2004)
251. T. Leismann, L. Anton, M.A. Aloy, E. Müller, J.M. Martí, J.A. Miralles, J.M. Ibáñez: *Relativistic MHD simulations of extragalactic jets*, A&A **436**, 503 (2005)
252. C.D. Levermore, G.C. Pomraning: *A flux-limited diffusion theory*, ApJ **248**, 321 (1981)
253. L.-X. Li, E.R. Zimmerman, R. Narayan, J.E. McClintock: *Multi-temperature blackbody spectrum of a thin accretion disk around a Kerr black hole: Model computations and comparison with observations*, ApJS **157**, 335 (2005)
254. E.P.T. Liang, R.H. Price: *Accretion disk coronae and Cygnus X-1*, ApJ **218**, 247 (1977)
255. J. Liebert, P. Bergeron, J.B. Holberg: *The formation rate, mass and luminosity functions of DA white dwarfs from the Palomar Green Survey*, ApJSup **156**, 47 (2005)
256. S. Liu, V. Petrosian, F. Melia: *Electron acceleration around the supermassive black hole at the Galactic Center*, ApJ **611**, L101 (2004)
257. F.S.N. Lobo: *Stable dark energy stars*, Class. Quant. Grav. **23**, 1525 (2006); gr-qc/0508115
258. O. Löhmer, W. Lewandowski, A. Wolszczan, R. Wielebinski: *Shapiro delay in the PSR J1640+2224 binary system*, ApJ **621**, 388 (2005); arXiv:astro-ph/0411742
259. C.P. Lorenz, D.G. Ravenhall, C.J. Pethick: *Neutron star crusts*, Phys. Rev. Lett. **70**, 379 (1993)
260. J.-P. Luminet: *Black holes: A general introduction*, in *Black Holes: Theory and Observation*, eds. F. Hehl, C. Kiefer, R. Metzler, Lecture Notes in Physics, p. 3 (Springer-Verlag, Berlin 1998)
261. A. Lyne, M. Burgay, M. Kramer et al.: *A double-pulsar system: A rare laboratory for relativistic gravity and plasma physics*, Science **303**, 1153 (2004); arXiv:astro-ph/0401086
262. D.A. MacDonald: *Numerical models of force-free black-hole magnetospheres*, MNRAS **211**, 313 (1984)
263. D.A. MacDonald, K.S. Thorne: *Black-hole electrodynamics – An absolute-space/universal-time formulation*, MNRAS **198**, 354 (1982)
264. P. MacNeice, K.M. Olson, C. Mobarry, R. deFainchtein, C. Packer: *PARAMESH: A parallel adaptive mesh refinement community toolkit*, Computer Physics Comm. **126**, 330 (2000)

265. J. Madej et al.: *Mass distribution of DA white dwarfs in the first data release of the SDSS*, A&A **419**, L5 (2004)
266. R.N. Manchester, G.B. Hobbs, A. Teoh, M. Hobbs: *The ATNF Pulsar Catalogue*, AJ **129**, 1993 (2005)
267. V.S. Manko, E.W. Mielke, W. Eckehard, J.D. Sanabria-Gómez: *Exact solution for the exterior field of a rotating neutron star*, Phys. Rev. D **61**, 081501 (2000)
268. V.S. Manko, J.D. Sanabria-Gómez, O.V. Manko: *Nine-parameter electrovac metric involving rational functions*, Phys. Rev. D **62**, 044048 (2000)
269. A. Marscher, S.G. Jorstad, J.L. Gomez et al.: *Observational evidence for the accretion-disk origin of a radio jet in an active galaxy*, Nature **617**, 625 (2002)
270. H.L. Marshall et al.: *A Chandra survey of quasar jets: First results*, ApJSupp **156**, 13 (2005)
271. J.M. Martí, E. Müller: *Numerical hydrodynamics in special relativity*, Living Reviews in Relativity **Irr-2003-6** (2003)
272. A. Martocchia, V. Karas, G. Matt: *Effects of Kerr space-time on spectral features from X-ray illuminated accretion discs*, MNRAS **312**, 817 (2000)
273. T. Maruyama, T. Tatsumi, D.N. Voskresensky, T. Tanigawa, S. Chiba: *Nuclear pasta structures and the charge screening effect*, Phys. Rev. C **72** 015802 (2005)
274. S. D. Mathur: *Where are the states of a black hole?*, arXiv:hep-th/0401115 (2004)
275. R. Matsumoto: *Three-dimensional global MHD simulations of accretion disks advection viscosity, and fluctuations*, Proc. of the Disk-Instability Workshop, 1998, Kyoto, Japan. Ed. by S. Mineshige and J. C. Wheeler. Frontiers Science Series No. 26 (Universal Academy Press, Inc. 1999), p 303
276. O. Maxwell, G.E. Brown, D.K. Campbell, R.F. Dashen, J.T. Manassah: *Beta decay of pion condensates as a cooling mechanism for neutron stars*, ApJ **216**, 77 (1977)
277. P. O. Mazur, E. Mottola: *Weyl cohomology and the effective action for conformal anomalies*, Phys. Rev. D **64**, 104022 (2001)
278. P. O. Mazur, E. Mottola: *Gravitational condensate stars*, arXiv:gr-qc/0109035 (2001)
279. P. O. Mazur, E. Mottola: *Dark energy and condensate stars: Casimir energy in the large*, arXiv:gr-qc/0405111 (2004)
280. P. O. Mazur, E. Mottola: *Gravitational vacuum condensate stars*, Proc. Nat. Acad. Sci. **111**, 9545 (2004); arXiv:gr-qc/0407075
281. J.C. McKinney: *General relativistic magnetohydrodynamic simulations of jet formation and large-scale propagation from black hole accretion systems*, MNRAS **368**, 1561 (2006); astro-ph/0603045
282. T.S. Metcalfe, M.H. Montgomery, A. Kanaan: *Testing white dwarf crystallization theory with asteroseismology of the massive pulsating DA Star BPM 37093*, ApJ **605**, L133 (2004); arXiv:astro-ph/0402046
283. J.E. McClintock, R.A. Remillard: *Black hole binaries*, in *Compact Stellar X-Ray Sources*, eds. W.H.G. Lewin and M. van der Klis (Cambridge Univ. Press, Cambridge 2004); arXiv:astro-ph/0306213
284. I.M. McHardy, I.E. Papadakis, P. Uttley: *Combined long and short time-scale X-ray variability of NGC 4051 with RXTE and XMM-Newton*, MNRAS **348**, 783 (2004)
285. J.C. McKinney: *Total and jet Blandford-Znajek power in presence of accretion disk*, ApJ **630**, L5 (2005); arXiv:astro-ph/0506367
286. J.C. McKinney: *Jet formation in black hole accretion systems I: Theoretical unification model*, arXiv:astro-ph/0506368 (2005)
287. J.C. McKinney: *Jet formation in black hole accretion systems II: Numerical models*, astro-ph/0506369 (2005)

288. J.C. McKinney: *A measurement of the electromagnetic luminosity of a Kerr black hole*, ApJ **611**, 977 (2004)
289. F. Melia, R. Coker: *Stellar gas flows into a dark cluster potential at the Galactic Center*, ApJ **511**, 750 (1999)
290. D. Merritt, M. Milosavljevic: *Massive binary black hole evolution*, Living Reviews in Relativity **8**, no. 8; arXiv:astro-ph/0410364 (2004)
291. A. Mignone, G. Bodo: *An HLLC Riemann solver for relativistic flows: I – Hydrodynamics*, MNRAS **364**, 126 (2005)
292. A. Mignone, G. Bodo: *An HLLC Riemann solver for relativistic flows: II – Magnetohydrodynamics*, MNRAS, in press (2006)
293. D. Mihalas, B.W. Mihalas: *Foundations of Radiation Hydrodynamics* (Oxford University Press, Oxford 1984)
294. C. Miller, F.K. Lamb, D. Psaltis: *Sonic-point model of kilohertz quasi-periodic brightness oscillations in low-mass X-ray binaries*, ApJ **508**, 791 (1998)
295. M.C. Miller, E.J.M. Colbert: *Intermediate-mass black holes*, Int. J. Mod. Phys. D **13**, 1 (2004); arXiv:astro-ph/0308402
296. F. Mirabel, L.F. Rodriguez: *A superluminal source in the Galaxy*, Nature **371**, 46 (1994)
297. F. Mirabel, L.F. Rodriguez: *Microquasars in our Galaxy*, Nature **392**, 673 (1998)
298. J. Miralda-Escudè, A. Gould: *A cluster of black holes at the Galactic Center*, ApJ **545**, 847 (2000)
299. R. Moderski, M. Sikora, J.-P. Lasota: *On the spin paradigm and the radio dichotomy of quasars*, MNRAS **301**, 142 (1998)
300. P. Möller, W.D. Myers, W.J. Swiatecki, J. Treiner, J.: *Nuclear mass formula with a finite-range droplet model and a folded-Yukawa single-particle potential*, Atomic Data and Nuclear Data Tables **39**, 225 (1988)
301. S.M. Morsink, L. Stella: *Relativistic precession around rotating neutron stars: Effects due to frame dragging and stellar oblateness*, ApJ **513**, 827 (1999)
302. E. Mottola, P.O. Mazur: *Gravitational condensate stars: An alternative to black holes*, APS, APRI2011M (2002)
303. A. Müller, M. Camenzind: *Relativistic emission lines from accreting black holes. The effect of disk truncation on line profiles*, A&A **413**, 861 (2004)
304. A. Müller: *Black hole astrophysics: Magnetohydrodynamics on the Kerr geometry*, PhD thesis, University of Heidelberg 2004
305. Müller, B.D. Serot: *Relativistic mean-field theory and the high-density nuclear equation of state*, Nucl. Phys. A **606**, 508 (1996)
306. M.P. Muno, E. Pfahl, F.K. Baganoff et al.: *An overabundance of transient X-ray binaries within 1 parsec of the Galactic Center*, ApJ **622**, L113 (2005)
307. R. Narayan, I. Yi: *Advection-dominated accretion: A self-similar solution*, ApJ **428**, L13 (1994)
308. R. Narayan, I. Yi: *Advection-dominated accretion: Underfed black holes and neutron stars*, ApJ **452**, 710 (1995)
309. R. Narayan, R. Mahadevan, J.E. Grindlay, R.G. Popham, C. Gammie: *Advection-dominated accretion model of Sagittarius A*: Evidence for a black hole at the Galactic center*, ApJ **494**, 554 (1998)
310. J.W. Negele, D. Vautherin: *Density-matrix expansion for an effective nuclear Hamiltonian. II*, Phys. Rev. C **11**, 1031 (1975)
311. S.C. Noble, C.F. Gammie, J.C. McKinney, L. Del Zanna: *Primitive variable solvers for conservative general relativistic magnetohydrodynamics*, astro-ph/0512420

312. T. Nozawa, N. Stergioulas, E. Gourgoulhon, Y. Eriguchi: *Construction of highly accurate models of rotating neutron stars – comparison of three different numerical methods*, A&ASuppl **132**, 431 (1998)
313. I. Okamoto, O. Kaburaki: *Thermodynamical and evolutional properties of Kerr black holes*, MNRAS **247**, 244 (1990)
314. I. Okamoto: *The evolution of a black hole's force-free magnetosphere*, MNRAS **254**, 192 (1992)
315. I. Okamoto: *Global asymptotic solutions for magnetohydrodynamic jets and winds*, ApJ **589**, 671 (2003)
316. T. Okuda, V. Teresi, E. Toscano, D. Molteni: *Black hole accretion discs and jets at super-Eddington luminosity*, MNRAS **357**, 295 (2005)
317. J.R. Oppenheimer, G.M. Volkoff: *On massive neutron cores*, Phys. Rev. **55**, 374 (1939)
318. J.R. Oppenheimer, H. Snyder: *On continued gravitational contraction*, Phys. Rev. **56**, 455 (1939)
319. B. Paczynsky, P.J. Wiita: *Thick accretion disks and supercritical luminosities*, A&A **88**, 23 (1980)
320. D.N. Page, K.S. Thorne: *Disk-accretion onto a black hole: I. Time-averaged structure of accretion disk*, ApJ **191**, 499 (1974)
321. D. Page, J.M. Lattimer, M. Prakash, A.W. Steiner: *Minimal cooling of neutron stars: A new paradigm*, ApJSupp **155**, 623 (2004)
322. C. Palenzuela-Luque, C. Bona: *Elements of numerical relativity*, Lecture Notes in Phys. **673** (Springer-Verlag, Berlin 2005)
323. V.R. Pandharipande, D.G. Ravenhall: *Hot nuclear matter*, in Proc. NATO Adv. Research Workshop on nuclear and heavy ion collisions, Les Houches, ed. M. Soyeur et al. (Plenum, New York 1989), p 103
324. J.A. Panei, L.G. Althaus, O.G. Benvenuto: *Mass–radius relations for white dwarf stars of different internal compositions*, A&A **353**, 970 (2000)
325. P. Papadopoulos, J.A. Font: *Relativistic hydrodynamics around black holes and horizon adapted coordinate systems*, Phys. Rev. D **58**, 24005 (1998)
326. P. Papadopoulos, J.A. Font: *Relativistic hydrodynamics on space-like and null surfaces: Formalism and computations of spherically symmetric spacetimes*, Phys. Rev. D **61**, 4015 (1998)
327. J. Peitz, S. Appl: *3+1 Formulation of non-ideal hydrodynamics*, MNRAS **296**, 231 (1998)
328. B.M. Peterson, L. Ferrarese, K.M. Gilbert et al.: *Central masses and broad-line region sizes of active galactic nuclei. II. A homogeneous analysis of a large reverberation-mapping database* ApJ **613**, 682 (2004)
329. Tsvi Piran: *The physics of gamma-ray bursts*, Rev. Mod. Phys. **76**, 1143 (2004)
330. J.A. Pons, J.M. Martí, E. Müller: *The exact solution of the Riemann problem with non-zero tangential velocities in relativistic hydrodynamics*, J. Fluid Mech. **422**, 125 (2000)
331. K. Powell, P.L. Roe, T.J. Linde et al.: *A solution-adaptive upwind scheme for ideal magnetohydrodynamics*, J. Comp. Phys. **154**, 284 (1999)
332. K.H. Prendergast, G.R. Burbidge: *On the nature of some Galactic X-ray sources*, ApJ **151**, L83 (1968)
333. J.L. Provencal, H.L. Shipman, E. Hog, P. Thejll: *Testing the white dwarf mass–radius relation with Hipparcos*, ApJ **494**, 759 (1998)
334. J.L. Provencal, H.L. Shipman, D. Koester, F. Wesemael, P. Bergeron: *Procyon B: Outside the iron box*, ApJ **568**, 324 (2002)

335. B. Punsly, F.V. Coroniti, V. Ferdinand: *Relativistic winds from pulsar and black hole magnetospheres*, ApJ **350**, 518 (1990)
336. B. Punsly, F.V. Coroniti: *Ergosphere-driven winds*, ApJ **354**, 583 (1990)
337. B. Punsly: *Black Hole Gravitohydromagnetics* (Springer-Verlag, Berlin 2001)
338. D.G. Ravenhall, C.J. Pethick: *Matter at large neutron excess and the physics of neutron star crusts*, Ann. Rev. Nucl. Part. Sci. **45**, 429 (1995)
339. S. Reddy: *Novel phases at high density and their roles in the structure and evolution of neutron stars*, Acta Phys. Polon. **B33**, 4101 (2002); arXiv:nucl-th/0211045
340. R.A. Remillard: *X-ray states of black hole binaries in outburst*, in *Interacting Binaries: Accretion, Evolution and Outcomes*, eds. L.A. Antonelli et al.; arXiv:astro-ph/0504126 (2005)
341. C.S. Reynolds, M.A. Nowak: *Fluorescent iron lines as a probe of astrophysical black hole systems*, Phys. Rep. **377**, 389 (2003)
342. H. Riffert, H. Herold: *Relativistic accretion disk structure revisited*, ApJ **450**, 508 (1995)
343. D.H. Rischke: *The quark-gluon plasma in equilibrium*, Progr. Particle Nucl. Phys. **52**, 197 (2004)
344. H. Ritter, U. Kolb: *Catalogue of cataclysmic binaries, low-mass X-ray binaries and related objects*, A&AS **129**, 83 (1998)
345. D.C. Robinson: *Uniqueness of the Kerr black hole*, Phys. Rev. Lett. **34**, 901 (1975)
346. P.L. Roe: *Approximate Riemann solvers, parameter vectors and difference schemes*, J. Comput. Phys. **43**, 357 (1981)
347. P.W.A. Roming, D. Vanden Berk, V. Palshin et al.: *GRB 060313: A new paradigm for short-hard bursts?*, astro-ph/0605005 (2006)
348. S.B. Ruster, V. Werth, M. Buballa, I.A. Shovkovy, D.H. Rischke: *The phase diagram of neutral quark matter: Self-consistent treatment of quark masses*, Phys. Rev. D **72**, 034004 (2005); arXiv:hep-ph/0503184
349. S.B. Ruster, V. Werth, M. Buballa, I.A. Shovkovy, D.H. Rischke: *The phase diagram of neutral quark matter: The effect of neutrino trapping*, Phys. Rev. D **73**, 034025 (2006); arXiv:hep-ph/0509073
350. S.B. Ruster, V. Werth, M. Buballa, I.A. Shovkovy, D.H. Rischke: *Phase diagram of neutral quark matter at moderate densities*, arXiv:nucl-th/0602018
351. D. Ryu, T.W. Jones, A. Frank: *Numerical magnetohydrodynamics in astrophysics: Algorithm and tests for multidimensional flow*, ApJ **452**, 785 (1995)
352. M. Salgado, S. Bonazzola, E. Gourgoulhon, P. Haensel: *High precision rotating neutron star models. I. Analysis of neutron star properties*, A&A **291**, 155 (1994)
353. M. Salgado, S. Bonazzola, E. Gourgoulhon, P. Haensel: *High precision rotating neutron star models. II. Large sample of neutron star properties*, A&ASup **108**, 455 (1994)
354. E.E. Salpeter: *Energy and pressure of a zero-temperature plasma*, ApJ **134**, 669 (1961)
355. K. Schertler, P.K. Sahu, C. Greiner, M.H. Thoma: *The influence of medium effects on the gross structure of hybrid stars*, Nucl. Phys. A **637**, 451 (1998)
356. K. Schertler, C. Greiner, J. Schaffner-Bielich, M.H. Thoma: *Quark phases in neutron stars and a third family of compact stars as a signature of phase transitions*, Nucl. Phys. A **677**, 463 (2000)
357. J.D. Schnittman, J.H. Krolik, J.F. Hawley: *Light curves from an MHD simulation of a black hole accretion disk*, astro-ph/0606615
358. K. Schöbel, M. Ansorg: *Maximal mass of uniformly rotating homogeneous stars in Einsteinian gravity*, A&A **405**, 405 (2003); arXiv:astro-ph/0301618
359. R. Schödel, T. Ott, R. Genzel, A. Eckart, N. Mouawad, T. Alexander: *Stellar dynamics in the central arcsecond of our Galaxy*, ApJ **596**, 1015 (2003)

360. K. Schwarzschild: Berliner Sitzungsberichte (Phys. Math. Klasse), 189–196, 3. Febr. (Mitt. 13. Jan. 1916) [Einstein himself wrote about this work: *Ihre Arbeit habe ich mit größtem Interesse durchgelesen. Ich hätte nicht erwartet, daß man so einfach die strenge Lösung der Aufgabe formulieren konnte. Die rechnerische Behandlung gefällt mir ausgezeichnet.* Shortly after this publication, Schwarzschild died on May 11, 1916.]
361. L. Segretain, G. Chabrier et al.: *Cooling theory of crystallized white dwarfs*, ApJ **434**, 641 (1994)
362. B.D. Serot, J.D. Walecka: *Properties of finite nuclei in a relativistic quantum field theory*, Phys. Lett. B **87**, 172 (1979)
363. N.J. Shakura, R. Sunyaev: *Black holes in binary systems. Observational appearance*, A&A **24**, 337 (1973)
364. I.I. Shapiro: *Fourth test of general relativity*, Phys. Rev. Letters **13**, 789 (1964)
365. N.J. Shaviv: *The nature of the radiative hydrodynamic instabilities in radiatively supported Thomson atmospheres*, ApJ **549**, 1093 (2001)
366. M. Shibata, T.W. Baumgarte, S.L. Shapiro: *The bar-mode instability in differentially rotating neutron stars: Simulations in full general relativity*, ApJ **542**, 453 (2000)
367. M. Shibata, Yu-ichiro Sekiguchi: *Gravitational waves from axisymmetric rotating stellar core collapse to a neutron star in full general relativity*, Phys. Rev. D **69**, 084024 (2004)
368. I.A. Shovkova: *Two lectures on color superconductivity*, arXiv:nucl-th/0410091 (2004)
369. C.-W. Shu, S. Osher: *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, J. Comp. Phys. **77**, 439 (1988)
370. N.R. Sibgatullin, N.M. Queen: *Oscillations and Waves in Strong Gravitational and Electromagnetic Fields* (Springer-Verlag, Berlin 1991)
371. N.R. Sibgatullin, R.A. Sunyaev: *Disk accretion in gravitational field of a rapidly rotating neutron star with a rotationally induced quadrupole mass moment*, Astron. Letters, **24**, 774 (1998)
372. S. Spindeldreher: *The discontinuous Galerkin method applied to the equations of ideal relativistic hydrodynamics*, PhD thesis, University of Heidelberg 2002
373. I.H. Stairs, S.E. Thorsett, J.H. Taylor, A. Wolszczan: *Studies of the relativistic binary pulsar PSR B1534+12: I. Timing analysis*, ApJ **616**, 414 (2004)
374. L. Stella, M. Vietri: *kHz quasiperiodic oscillations in low-mass X-Ray binaries as probes of general relativity in the strong-field regime*, Phys. Rev. Lett. **82**, 17 (1999)
375. N. Stergioulas: *The structure and stability of rotating relativistic stars*, PhD thesis, University of Wisconsin-Milwaukee, Milwaukee, USA, 1996
376. N. Stergioulas: *Rotating stars in relativity*, Living Reviews in Relativity **lrr-2003-3** (2003)
377. N. Stergioulas, J.L. Friedman: *Comparing models of rapidly rotating relativistic stars constructed by two numerical methods*, ApJ **444** 306 (1995)
378. J. Stone, D. Mihalas, M.L. Norman: *ZEUS-2D: A radiation magnetohydrodynamics code for astrophysical flows in two space dimensions. III – The radiation hydrodynamic algorithms and tests*, ApJSupp **80**, 819 (1992)
379. T.E. Strohmayer, C.B. Markwardt: *On the frequency evolution of X-ray brightness oscillations during thermonuclear X-ray bursts: Evidence of coherent oscillations*, ApJ **516**, L81 (1999)
380. T.E. Strohmayer, L. Bildsten: *New views of thermonuclear bursts*, arXiv:astro-ph/0301544
381. T.E. Strohmayer: *Future probes of the neutron star equation of state using X-ray bursts*, In: X-ray Timing 2003: Rossi and Beyond. AIP Conf. Proc. **714**, Ed. P. Kaaret, F.K. Lamb, J.H. Swank, Am. Inst. Phys. (2004), p. 245; arXiv:astro-ph/0401465

382. M. Stute, M. Camenzind: *Towards a self-consistent relativistic model of the exterior gravitational field of rapidly rotating neutron stars*, MNRAS **336**, 831 (2002)
383. K. Sumiyoshi, J.M. Ibanez, J.V. Romero: *Thermal history and structure of rotating protoneutron stars with relativistic equation of state*, A&ASupp **134**, 39 (1999)
384. L. Susskind: *The world as a hologram*, arXiv:hep-th/9409089 (1994)
385. M. Takahashi: *Transmagnetosonic accretion in a black hole magnetosphere*, ApJ **570**, 264 (2002)
386. Y. Tanaka, K. Nandra, A.C. Fabian et al.: *Gravitationally redshifted emission implying an accretion disk and massive black-hole in the active galaxy MCG:-6-30-15*, Nature **375**, 659 (1995)
387. K. Taniguchi, T.W. Baumgarte, J.A. Faber, S.L. Shapiro: *Black hole-neutron star binaries in general relativity: Effects of neutron star spin*, Phys. Rev. D **72**, 4008 (2005); arXiv:astro-ph/0505450
388. M. Tassoul, G. Fontaine, D.E. Winget: *Evolutionary models for pulsation studies of white dwarfs*, ApJSupp **72**, 335 (1990)
389. J. Taylor, J.M. Weisberg: *Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16*, ApJ **345**, 434 (1989)
390. A. Timokhin: *High resolution numerical modeling of the force-free pulsar magnetosphere*, AIP Conf. Proc. **801**, 332 (2005); arXiv:astro-ph/0507054 (2005)
391. G. 't Hooft: *The giant leap to the Planck length*, In: Perspectives on High Energy Physics and Cosmology, Eds. A. Gonzalez-Arroyo and C. Lopez (World Scientific, Singapore 1993), pp. 1-9
392. K.S. Thorne, R.H. Price, a. MacDonald: *Black holes: The membrane paradigm* (Yale University Press, Yale 1986)
393. S.E. Thorsett, D. Chakrabarty: *Neutron star masses measurements. I. Radio pulsars*, ApJ **512**, 288 (1999)
394. R.C. Tolman: *Static solution of Einstein's equations for spheres of fluids*, Phys. Rev. **55**, 364 (1939)
395. A. Tomimatsu: *Relativistic dynamos in magnetospheres of rotating compact objects*, ApJ **528**, 972 (2000)
396. E.F. Toro: *Riemann solvers and numerical methods for fluid dynamics*, 2nd Edn (Springer-Verlag, Berlin 1999).
397. S. Torres, E. Garcia-Berro, J. Isern, F. Figueras: *Simulating Gaia performances on white dwarfs*, MNRAS **360**, 1381 (2005)
398. Toth: *The $\nabla \cdot \mathbf{B} = 0$ Constraint in shock-capturing magnetohydrodynamics codes*, J. Comp. Phys. **161**, 605 (2000)
399. S. Tremaine, K. Gebhardt, R. Bender et al.: *The Slope of the black hole mass versus velocity dispersion correlation*, ApJ **574**, 740 (2002)
400. J. Trümper, V. Burwitz, F. Haberl, E. Zavlin: *The puzzles of RX J1856.5-3754: Neutron star or quark star?*, Nucl. Phys. B Proc. Suppl. **132**, 560 (2004)
401. N.J. Turner, J.M. Stone: *A module for radiation hydrodynamic calculations with ZEUS-2D using flux-limited diffusion*, ApJSupp **135**, 95 (2001)
402. N.J. Turner, J.M. Stone, J.H. Krolik, T. Sano: *Local three-dimensional simulations of magnetorotational instability in radiation-dominated accretion disks*, ApJ **593**, 992 (2003)
403. N.J. Turner, O.M. Blaes, A. Socrates et al.: *The effects of photon bubble instability in radiation-dominated accretion disks*, ApJ **624**, 267 (2005); arXiv:astro-ph/0501198
404. Y. Uchiyama, C.M. Urry, C.C. Cheung et al.: *Shedding new light on the 3C 273 jet with the Spitzer Space Telescope*, astro-ph/0605530

405. M.H.P.M. van Putten: J. Comp. Phys. **99**, 341 (1993)
406. M.H.P.M. van Putten, A. Levinson: *Theory and astrophysical consequences of a magnetized torus around a rapidly rotating black hole*, ApJ **584**, 937 (2003)
407. E.P. Velikov: Sov. Phys. JETP **36**, 995 (1959)
408. R.V. Wagoner: *Test for the existence of gravitational radiation*, ApJ **196**, 63 (1975)
409. R.M. Wald: *Black hole in a uniform magnetic field*, Phys. Rev. D **10**, 1680 (1974)
410. J.D. Walecka: *A theory of highly condensed matter*, Ann. Phys. **83**, 491 (1974)
411. M. Walker, R. Penrose: *On quadratic first integrals of the geodesic equations for type {22} spacetimes*, Comm. Math. Phys. **18**, 265 (1970)
412. D.-X. Wang, W.-H. Lei, Kan Xiao, Ren-Yi Ma: *An analytical model of black hole evolution and gamma ray bursts*, ApJ **580**, 358 (2002); arXiv:astro-ph/0209579
413. G. Watanabe, K. Sato, K. Yasuoka, T. Ebisuzaki: *Phases of hot nuclear matter at subnuclear densities*, Phys. Rev. C **69**, 055805 (2004)
414. G. Watanabe, T. Maruyama, K. Sato, K. Yasuoka, T. Ebisuzaki: *Simulation of transitions between “pasta” phases in dense matter*, Phys. Rev. Lett. **94**, 031101 (2005)
415. J.M. Weisberg, J. Taylor: *Relativistic binary pulsar B1913+16: Thirty years of observations and analysis*, arXiv:astro-ph/0407149 (2004)
416. M. Visser, D.L. Wiltshire: *Stable gravastars – An alternative to black holes?*, Class. Quant. Grav. **21** 1135 (2004); arXiv:gr-qc/0310107
417. F. Weber: *Strange quark matter and compact stars*, arXiv:astro-ph/0407155 (2004)
418. D.T. Wickramasinghe, L. Ferrario: *Magnetism in isolated white dwarf stars*, PASP **112**, 873 (2000)
419. C. Will: *The confrontation between general relativity and experiment*, Living Reviews in Relativity **Irr-2001-4** (2001)
420. J.R. Wilson: *Models of differentially rotating stars*, ApJ **176**, 195 (1972)
421. D.E. Winget, C.J. Hansen, J. Liebert et al.: *An independent method for determining the age of the Universe*, ApJ **315**, 77 (1987)
422. M.A. Wood: *Astro-archaeology: Reading the Galactic history recorded in the white dwarf stars*, PhD thesis, University of Texas at Austin 1990
423. P.R. Woodward, P. Collella: *The Piecewise Parabolic Method (PPM) for gas-dynamical simulations*, J. Comp. Phys. **54**, 115 (1984)
424. D.G. Yakovlev, C.J. Pethick: *Neutron star cooling*, ARA&A **42**, 169 (2004)
425. D.G. Yakovlev, O.Y. Gnedin, M.E. Gusakov et al.: *Neutron star cooling*, Nucl. Phys. A **752**, 590 (2005)
426. M. Yokosawa: *Energy and angular momentum transport in magnetohydrodynamical accretion onto a rotating black hole*, PASJ **45**, 207 (1993)
427. M. Yokosawa: *Structure and dynamics of an accretion disk around a black hole*, PASJ **47**, 605 (1995)
428. M. Yokosawa, T. Inui: *Magnetorotational instability around a rotating black hole*, ApJ **631**, 1051 (2005); arXiv:astro-ph/0503712
429. J. York: *Kinematics and dynamics of general relativity*, in *Sources of Gravitational Radiation*, ed. L.L. Smarr (Cambridge University Press, Cambridge 1979)
430. Feng Yuan, Wei Cui, R. Narayan: *An accretion-jet model for black hole binaries: Interpreting the spectral and timing features of XTE J1118+480*, ApJ **620**, 905 (2005)
431. U. Ziegler: *NIRVANA+: An adaptive mesh refinement code for gas dynamics and MHD*, Comp. Phys. Comm. **109**, 142 (1998)
432. U. Ziegler: *A three-dimensional Cartesian adaptive mesh code for compressible magnetohydrodynamics*, Comp. Phys. Comm. **116**, 65 (1999)

Index

40 Eri B 162

Accretion

- 1D models 542
- ADAF models 551
- angular momentum transport 514
- cooling 544
- GRMHD 559
- jets and ergosphere 566
- magnetorotational instability 514
- MRI simulations 528
- photon bubbles 536, 537
- quasiperiodic oscillations 533
- radiation-pressure dominated 536
- radiative MHD 517
- spectral energy distribution 549
- super-Eddington 553
- time-scales 556
- truncated 540
- turbulent angular momentum transport 539
- turbulent states 538
- two-temperature plasma 534

Accretion torus 529

- relativistic 563

AM Her systems 183

AMR

- adaptive mesh refinement 523

Angular momentum transport

- disks 513
- Rayleigh criterium 514

Bardeen observer 309

Binary system

- Cyg X-1 452
- double pulsar 265
- gravitational radiation 251
- mass function 451

merger time 253

microquasars 453

post-Newtonian effects 244

PSR B1913+16 262

pulsar timing 256

Römer time delay 259

X-ray system 452

Black hole

- Doppler factor 438
- Eddington–Finkelstein 359
- Einstein 4
- entropy 403
- equatorial geodesics 416
- ergosphere 389
- event horizon 386
- Fe K line 439
- first law 403
- four laws 400
- Galactic center 461
- geodesics 4th integral 424
- geodesics Killing tensor 424
- holographic principle 404
- ISCO 392
- Kerr solution 379
- Killing horizon 388
- Kruskal extension 359
- magnetosphere 487
- marginally stable radius 392
- mass-plane 450
- merging 411
- observed emission 438
- Penrose diagram 363
- quasars 468
- ray-tracing 432
- rotational energy 405
- Salpeter time 410
- Schwarzschild 356
- second law 407

- Sgr A* 462
- spin evolution 409
- supermassive 16
- supermassive bulge correlation 465
- surface gravity 400
- temperature 403
- third law 408
- tortoise coordinates 357
- uniqueness theorem 385
- Weyl–Papapetrou form 384

- Chandrasekhar 4
- Chombo 523
- Compact objects 6
- Congruence of time-like geodesics 73
- Connection
 - axisymmetric 600
 - Bianchi identities 58
 - Cartan equations 61
 - Levi-Civita 52
 - Lorentzian 65
 - spin connection 64
 - torsion 57
- Cowling's theorem 484
- Crab Nebula
 - spectrum 620
- Curvature
 - axisymmetric hypersurfaces 602
 - axisymmetric spacetime 624
 - Ricci tensor 60
 - Riemann tensor 57
 - two-sphere 63
 - Weyl tensor 60
- Cygnus X-1
 - bimodal spectrum 11
 - Uhuru 9
 - X-ray emission 452
- DQ Her systems 184
- Dwarf novae 183

- Einstein's equations
 - 3+1 axisymmetric 320
 - 3+1 split 100
 - black holes 380
 - constraints equations 98
 - cosmological constant 69
 - Poisson equation 98
- Electron degeneracy 152

- EoS
 - above neutron drip 190
 - analytical fits 217
 - APR98 199
 - below neutron drip 153
 - BPAL21 199
 - BPS 192
 - degenerate electrons 156
 - electrostatic corrections 157
 - gravastar 444
 - low densities 158
 - mean field theory 207
 - quark matter 237
 - SLy4 199
- Ernst potentials 343
- Extrinsic curvature
 - axisymmetric 600
- Extrinsic curvature, definition 93
- Extrinsic curvature, Lie derivative 95

- Fermi gas 154

- Galactic center
 - stellar black holes 464
- Geodetic deviation 73
- Geodesics
 - Hamilton–Jacobi equation 425
- GR
 - Einstein equivalence principle 29
 - quantum theory 72
 - strong equivalence principle 34
 - weak equivalence principle 29
- GR Hydro equations
 - plasma equation 106
- Gravastar
 - anisotropic pressure 445
 - EoS 444
 - essentials 443
 - geodesics 630
 - metric element 443
- Gravitational Radiation
 - quadrupole formula 251
- Gravitational redshift 31
- Gravity
 - metric theories 33
- Gravity Probe B 312
- GRBs
 - afterglow 19
 - detection 19

- GRHydro
 - mechanical equilibrium 625
- GRMHD
 - energy-momentum tensor 606
 - conservative formulation 607
 - constrained flux transport 612
 - fluxes 561
 - HARMS code 563
 - jet formation 559
 - Maxwell's equations 607
 - primitive variables 609
- Hawking radiation 403
- Hilbert action 68
- Instability
 - magnetorotational 514
 - weak magnetic field 515
- Inverse beta-decay 159
- Isospin current 214
- Kerr solution 378
 - Boyer–Lindquist coordinates 379
 - Carter integrals 416
 - derivation 380
 - horizon angular velocity 389
 - Kerr–Schild form 398
 - Penrose diagram 394
 - rotational energy 400
 - surface gravity 401
 - uniqueness 385
- Kruskal metric
 - Penrose diagram 366
- LAGEOS satellites 314
- LMXB
 - binary system 297
- Loop quantum gravity 72
- Lorene 337
- Magnetosphere
 - Goldreich–Julian 491
 - Grad–Shafranov equation 489
 - jet formation 493
 - light cylinder 490
 - light cylinder function 490
 - null surface 491
- Manifold
 - parallel transport 50
- affine connection 47
- covariant derivative 47
- divergence of vector fields 55
- exterior derivative 45
- one-forms 39
- tangent vectors 37
- tensors 40
- vector bundle 42
- vector fields 41
- vielbein 38
- volume element 46
- Manko solution 345
- Mass–radius
 - hadronic stars 223
- Maxwell's equations
 - Current flux 482
 - Grad–Shafranov 480
- Messier 87
 - radio image 568
- Messier87
 - jets 568
- MHD equations
 - advective 517
 - curvilinear 519
 - conservative 518
 - divergence cleaning 525
 - HLL flux 522
 - hyperbolic 521
 - radiation 536
 - state vector 519
 - two-temperature 534
- Microquasars
 - comparison with quasars 454
 - properties 453
- Minkowski space
 - Penrose diagram 363
- MIT bag model 228
- MRI
 - dispersion relation 515
 - existence 514
 - global simulations 529
 - growth condition 527
 - quasiperiodic oscillations 533
- Neutron star
 - binding energy 205
 - accretion power 297
 - color superconductivity 226
 - cooling 276

- cooling curve 281
- core collapse 351
- flavors 189
- galactic 270
- hadronic models 214, 221
- heat capacity 279
- internal structure 188
- mass-radius relation 203
- masses 244
- massive 242
- mixed phases 241
- moment of inertia 206
- neutrino emission 277
- pion condensate 277
- QPOs 300
- quark core 237
- rotation powered 270
- superconductivity 225
- surface redshift 205
- thermal emission 272
- TOV equations 201

- Observer**
 - Eulerian 93, 108, 559
 - Newtonian 28
 - Relativistic 43
 - Tetrad 65
 - ZAMO 476

- Paramesh 523
- Penrose diagram 366
 - conformal structure 363
 - Kruskal 368
 - Minkowski 364
- Periastron shift 248
 - PSR B1913+16 250
- Poisson equation
 - black holes 381
 - rotating star 322
- Post-Newtonian
 - gauge freedom 245
 - Lagrangian point particles 248
 - potentials 245
 - Shapiro time delay 250
- Precession
 - geodetic 312
 - Lense–Thirring 315
- Procyon B 162
- PSR B1534+12

- parameters 262
- PSR B1913+16 250
 - energy loss 253
 - orbit decay 253
 - parameters 262
- PSR J0737–3039
 - A+B 265
 - masses 268

- QPO**
 - magnetorotational 533
- Quark matter**
 - CFL 232
 - gapless CFL 233
 - grand canonical ensemble 231
- Quark star**
 - strange 241
- Quark stars**
 - phase diagram 224
- Quasar**
 - luminosity 468

- Radio pulsar**
 - breaking index 284
 - gamma-ray emission 290
 - slow-down 284
 - X-ray emission 290
- Radio pulsars**
 - double pulsar system 265
 - global spectra 287
 - masses of companion 264
 - millisecond 14
 - table 264
 - times of arrival TOAs 259
 - timing 255
 - timing formula 260
- Ray-tracing**
 - Fe lines 442
 - Kerr geometry 432
 - photon momenta 437
- Ricci tensor**
 - 4D axisymmetric 319
 - axisymmetric surfaces 602
 - definition 60
- Riemann tensor**
 - axisymmetric spacetimes 624
 - Bianchi identities 59
 - Kerr metric 629
 - local coordinates 63

- Schwarzschild 126
- symmetries 59
- two-sphere 64
- Rotating star
 - bar mode instability 351
 - binding energy 331
 - hydrostatic equilibrium 324
 - Manko solution 345
 - mass–radius 340
 - masses 326
 - metric element 321
 - numerical integration 335
 - oscillations 350
 - structure equations 322
- Rotation
 - differential 311
- Schwarzschild solution
 - black hole 356
 - derivation 130
 - effective potential 371
 - Kruskal extension 359
 - orbital equation 373
 - perihelion advance 378
- Sigma–omega model 209
- Sirius B
 - temperature 138
 - white dwarf 138
- Spacetime
 - 3+1 split 316
 - spherically symmetric 123
 - axisymmetric 308
 - axisymmetric isotropic 321
 - Hartle–Thorne metric 333
 - Kerr solution 378
 - manifold 42
 - metric tensor 43
 - particle motion 54
 - Weyl–Papapetrou 342
- Stein 2051 162
- Stellar structure
 - TOV equations 124
 - Supersoft sources 182
 - Transverse acceleration 74
 - Turbulence
 - magnetic 513
 - Velocity field
 - local observers 310
 - Virial theorem
 - white dwarfs 169
 - Weyl–Papapetrou metric 342
 - Whisky code 352
 - White dwarfs
 - cataclysmic systems 180
 - classical novae 181
 - diamond 151
 - globular cluster 143
 - instability 167
 - age of Universe 149
 - binary systems 180
 - central density 151
 - chronometers 146
 - cooling curves 147, 178
 - disk age 148
 - GAIA luminosity function 149
 - GR instability 171
 - Hipparcos data 162
 - intermediate polars 184
 - magnetic 144
 - mass distribution 142
 - mass–radius relation 161
 - polars 183
 - polytropic EoS 160
 - SDSS 141
 - spectral classification 139
 - structure equations 160
 - surface layers 175
 - total energy 169



ASTRONOMY AND ASTROPHYSICS LIBRARY

Series Editors:

G. Börner · A. Burkert · W. B. Burton · M. A. Dopita
A. Eckart · T. Encrenaz · E. K. Grebel · B. Leibundgut
J. Lequeux · A. Maeder · V. Trimble

The Stars By E. L. Schatzman and F. Praderie

Modern Astrometry 2nd Edition

By J. Kovalevsky

The Physics and Dynamics of Planetary Nebulae By G. A. Gurzadyan

Galaxies and Cosmology By F. Combes, P. Boissé, A. Mazure and A. Blanchard

Observational Astrophysics 2nd Edition
By P. Léna, F. Lebrun and F. Mignard

Physics of Planetary Rings Celestial Mechanics of Continuous Media
By A. M. Fridman and N. N. Gorkavyi

Tools of Radio Astronomy 4th Edition,
Corr. 2nd printing
By K. Rohlfs and T. L. Wilson

Tools of Radio Astronomy Problems and Solutions 1st Edition, Corr. 2nd printing
By T. L. Wilson and S. Hüttemeister

Astrophysical Formulae 3rd Edition (2 volumes)
Volume I: Radiation, Gas Processes and High Energy Astrophysics
Volume II: Space, Time, Matter and Cosmology
By K. R. Lang

Galaxy Formation By M. S. Longair

Astrophysical Concepts 4th Edition
By M. Harwit

Astrometry of Fundamental Catalogues
The Evolution from Optical to Radio Reference Frames
By H. G. Walter and O. J. Sovers

Compact Stars. Nuclear Physics, Particle Physics and General Relativity 2nd Edition
By N. K. Glendenning

The Sun from Space By K. R. Lang

Stellar Physics (2 volumes)
Volume 1: Fundamental Concepts and Stellar Equilibrium
By G. S. Bisnovatyi-Kogan

Stellar Physics (2 volumes)

Volume 2: Stellar Evolution and Stability
By G. S. Bisnovatyi-Kogan

Theory of Orbits (2 volumes)
Volume 1: Integrable Systems and Non-perturbative Methods
Volume 2: Perturbative and Geometrical Methods
By D. Boccaletti and G. Pucacco

Black Hole Gravitohydromagnetics
By B. Punsly

Stellar Structure and Evolution
By R. Kippenhahn and A. Weigert

Gravitational Lenses By P. Schneider, J. Ehlers and E. E. Falco

Reflecting Telescope Optics (2 volumes)
Volume I: Basic Design Theory and its Historical Development. 2nd Edition
Volume II: Manufacture, Testing, Alignment, Modern Techniques
By R. N. Wilson

Interplanetary Dust
By E. Grün, B. Å. S. Gustafson, S. Dermott and H. Fechtig (Eds.)

The Universe in Gamma Rays
By V. Schönfelder

Astrophysics. A New Approach 2nd Edition
By W. Kundt

Cosmic Ray Astrophysics
By R. Schlickeiser

Astrophysics of the Diffuse Universe
By M. A. Dopita and R. S. Sutherland

The Sun An Introduction. 2nd Edition
By M. Stix

Order and Chaos in Dynamical Astronomy
By G. J. Contopoulos

Astronomical Image and Data Analysis
2nd Edition By J.-L. Starck and F. Murtagh

The Early Universe Facts and Fiction
4th Edition By G. Börner



ASTRONOMY AND ASTROPHYSICS LIBRARY

Series Editors:

G. Börner · A. Burkert · W. B. Burton · M. A. Dopita
A. Eckart · T. Encrenaz · E. K. Grebel · B. Leibundgut
J. Lequeux · A. Maeder · V. Trimble

The Design and Construction of Large Optical Telescopes By P. Y. Bely

The Solar System 4th Edition

By T. Encrenaz, J.-P. Bibring, M. Blanc,
M. A. Barucci, F. Roques, Ph. Zarka

General Relativity, Astrophysics, and Cosmology By A. K. Raychaudhuri,
S. Banerji, and A. Banerjee

Stellar Interiors Physical Principles,
Structure, and Evolution 2nd Edition
By C. J. Hansen, S. D. Kawaler, and V. Trimble

Asymptotic Giant Branch Stars
By H. J. Habing and H. Olofsson

The Interstellar Medium
By J. Lequeux

Methods of Celestial Mechanics (2 volumes)
Volume I: Physical, Mathematical, and
Numerical Principles
Volume II: Application to Planetary System,
Geodynamics and Satellite Geodesy
By G. Beutler

Solar-Type Activity in Main-Sequence Stars
By R. E. Gershberg

Relativistic Astrophysics and Cosmology
A Primer By P. Hoyng

Magneto-Fluid Dynamics
Fundamentals and Case Studies
By P. Lorrain

Compact Objects in Astrophysics
White Dwarfs, Neutron Stars and Black Holes
By Max Camenzind

Special and General Relativity
With Applications to White Dwarfs, Neutron
Stars and Black Holes
By Norman K. Glendenning